

Canonical Duality-Triality Theory: Bridge Between Nonconvex Analysis/Mechanics and Global Optimization in Complex Systems

David Yang Gao, Ning Ruan

Faculty of Sciences and Technology, Federation University Australia, Mt Helen, VIC 3353, Australia.

Vittorio Latorre

Department of Computer, Control and Management Engineering, Sapienza University of Rome, Rome, Italy

Dedicated to Professor Gilbert Strang on the Occasion of His 80th Birthday

Abstract

Canonical duality-triality is a breakthrough methodological theory, which can be used not only for modeling complex systems within a unified framework, but also for solving a wide class of challenging problems from real-world applications. This paper presents a brief review on this theory, its philosophical origin, physics foundation, and mathematical statements in both finite and infinite dimensional spaces. Particular emphasis is placed on its role for bridging the gap between nonconvex analysis/mechanics and global optimization. Special attentions are paid on unified understanding the fundamental difficulties in large deformation mechanics, bifurcation/chaos in nonlinear science, and the NP-hard problems in global optimization, as well as the theorems, methods, and algorithms for solving these challenging problems. Misunderstandings and confusion on some basic concepts, such as objectivity, nonlinearity, Lagrangian, and generalized convexities are discussed and classified. Breakthrough from recent challenges and conceptual mistakes by M. Voisei, C. Zălinescu and his co-worker are addressed. Some open problems and future works in global optimization and nonconvex mechanics are proposed.

Keywords: Duality, complementarity, triality, mathematical modeling, large deformation, nonlinear PDEs, NP-hard problems, nonconvex analysis, global optimization.

1 Introduction

Duality is one of the oldest and most beautiful ideas in human knowledge. It has a simple origin from the oriental philosophy of *yin-yang principle* tracing back 5000 years ago. According to *I Ching*¹, the fundamental law of the nature is the *Dao*, the duality of one yin and one yang, which gives two opposite or complementary points of view of looking at

¹Also known as the *Book of Changes*, *Zhouyi* and *Yijing*, is the worlds oldest and most sophisticated system of wisdom divination, the fundamental source of most of the easts philosophy, medicine and

the same object. In quantum mechanics, the wave-particle duality is a typical example to fully describe the behavior of quantum-scale objects. Mathematically, duality represents certain translation of concepts, theorems or mathematical structures in a one-to-one fashion, i.e., if the dual of A is B, then the dual of B is A (cf. [5, 21, 125]). This one-to-one complementary relation is called the *canonical duality*. It is emphasized recently by Sir Michael Atiyah that duality in mathematics is not a theorem, but a “principle” [5]. Therefore, any duality gap is not allowed. This fact is well-known in mathematics and physics, but not in optimization due to the existing gap between these fields. To bridge this gap, a canonical duality-triality theory has been developed originally from nonconvex mechanics [53] with extensive applications in engineering, mathematics, and sciences, especially in the multidisciplinary fields of nonconvex mechanics and global optimization [61, 69, 84].

1.1 Nonconvex analysis/mechanics and difficulties

Mathematical theory of duality for convex problems has been well-established. In linear elasticity, it is well-known that each potential energy principle is associated with a unique complementary energy principle through Legendre transformation. This one-to-one duality is guaranteed by convexity of the stored energy. The well-known Helinger-Reissner principle is actually a special Lagrangian saddle min-max duality theory in convex analysis, which lays a foundation for mixed/hybrid finite element methods with successful applications in structural limit analysis [32, 33]. However, the one-to-one duality is broken in nonconvex systems. In large deformation theory, the stored energy is generally nonconvex and its Legendre conjugate can't be uniquely determined. It turns out that the existence of a pure stress-based complementary-dual energy principle (no duality gap) was a well-known open problem over a half century and subjected to extensive discussions by many leading experts including Levison [112], Koiter [101], Oden and Reddy [128], Ogden [130], Lee and Shield [111], Stumpf [146], etc.

Nonconvex phenomena arise naturally in large classes of engineering applications. Many real-life problems in modern mechanics and complex systems require consideration of nonconvex effects for their accurate modelling. For example, in modelling of hysteresis, phase transitions, shape-memory alloys, and super-conducting materials, the free energy functions are usually nonconvex due to certain internal variables [70, 75, 76]. In large deformation analysis, thin-walled structure can buckle even before the stress reaches its elastic limit [41, 42, 86]. Mathematically speaking, many fundamentally difficult problems in engineering and the sciences are mainly due to the nonconvexity of their modelling. In static systems, the nonconvexity usually leads to multi-solutions in the related governing equations. Each of these solutions represents certain possible phase or buckled state in large deformed solids. These local solutions are very sensitive to the internal parameters and external force. In dynamical systems, the so-called chaotic behavior is mainly due to nonconvexity of the objective functions [60]. Numerical methods (such as FEM, FDM, etc) for solving nonconvex minimal potential variational problems usually end up with nonconvex optimization problems [44, 55, 92, 98, 143]. Due to the lack of global optimality criteria, finding global optimal solutions is fundamentally difficult, or even impossible by traditional numerical methods and optimization techniques. For example, it was discovered by Gao and Ogden [75, 76] that for certain given external

spirituality. Traditionally it was believed that the principles of the I Ching originated with the mythical King Fu Xi during the 3rd and 2nd millennia BCE.

loads, both the global and local minimizers are nonsmooth and cannot be determined by any Newton-type numerical methods. In fact, many nonconvex problems are considered as NP-hard (Non-deterministic Polynomial-time hard) in global optimization and computer science [69, 84]. Unfortunately, these well-known difficulties are not fully recognized in computational mechanics due to the significant gap between engineering mechanics and global optimization. Indeed, engineers and scientists are mistakenly attempting to use traditional finite element methods and commercial software for solving nonconvex mechanics problems. In order to identify the fundamental difficulty of the nonconvexity from the traditional definition of nonlinearity, the terminology of *Nonconvex Mechanics* was formally proposed by Gao, Ogden and Stavroulakis in 1999 [77]. The *Handbook of Nonconvex Analysis* by Gao and Motreanu [74] presents recent advances in the field.

1.2 Global optimization and challenges

In parallel with the nonconvex mechanics, global optimization is a multi-disciplinary research field developed mainly from nonconvex/combinatorial optimization and computational science during the last nineties. In general, the global optimization problem is formulated in terms of finding the absolutely best set of solutions for the following constrained optimization problem

$$\min f(x), \quad \text{s.t.} \quad h_i(x) = 0, \quad g_j(x) \leq 0 \quad \forall i \in I_m, \quad j \in I_p, \quad (1)$$

where $f(x)$ is the so-called “objective function”², $h_i(x)$ and $g_j(x)$ are constraint functions, $I_m = \{1, \dots, m\}$ and $I_p = \{1, \dots, p\}$ are index sets. It must be emphasized that, different from the basic concept of *objectivity* in continuum physics, the objective function extensively used in mathematical optimization is allowed to be any arbitrarily given function, even the linear function. Clearly, this mathematical model is artificial. Although it enables one to “model” a very wide range of problems, it comes at a price: even very special kinds of nonconvex/discrete optimization problems are considered to be NP-hard. This dilemma is due to the gap between mathematical optimization and mathematical physics. In science, the concept of objectivity is often attributed with the property of scientific measurements that can be measured independently of the observer. Therefore, a function in mathematical physics is called objective only if it depends on certain measure of its variables (see Definition 6.1.2, [53] and the next section). Generally speaking, a useful mathematical model must obey certain fundamental law of nature. Without detailed information on these arbitrarily given functions, it is impossible to have a general theory for finding global extrema of the general nonconvex problem (1). This could be the reason why there was no breakthrough in nonlinear programming during the past 60 years.

In addition to the nonconvexity, many global optimization problems in engineering design and operations research explicitly require integer or binary decision variables. For example, in topology optimization of engineering structures, the design variable of material density $\rho(\mathbf{x}) = \{0, 1\}$ is a discrete selection field, i.e. by selection it has to take the value, 1, and by de-selection it has to take the value, 0 (see [9]). By the fact that the deformation variable is a continuous field, which should be determined in each iteration for topological structure, therefore, the finite element method for solving topology

²This terminology is used mainly in English literature. The function $f(x)$ is called the target function in Chinese and Japanese literatures, the goal function in Russian and German literatures.

optimization problems ends up with a coupled mixed integer nonlinear programming problem. Discrete problems are frequently encountered in modeling real world systems for a wide spectrum of applications in decision science, management optimization, industrial and systems engineering. Imposing such integer constraints on the variables makes the global optimization problems much more difficult to solve. It is well-known in computational science and global optimization that even the most simple quadratic minimization problem with boolean constraint

$$\min \left\{ \frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} - \mathbf{x}^T \mathbf{f} \mid \mathbf{x} \in \{0, 1\}^n \right\} \quad (2)$$

is considered to be NP-hard (Non-deterministic Polynomial-time hard) [80]. Indeed, this integer minimization problem has 2^n local solutions. Due to the lack of global optimality criterion, traditional direct approaches, such as the popular branch and bound methods, can only handle very small size problems. Actually, it was proved by Pardalos and Vavasis [132, 150] that instead of the integer constraint, the continuous quadratic minimization with box constraints $\mathbf{x} \in [0, 1]^n$ is NP-hard as long as the matrix \mathbf{Q} has one negative eigenvalue.

During the last 20 years, the field of global optimization has been developed dramatically to across almost every branch of sciences, engineering, and complex systems [30, 31, 135]. By the fact that the mathematical model (1) is too general to have a mathematical theory for identifying global extrema, the main task in global optimization is to study algorithmic methods for numerically solving the optimal solutions. These methods can be categorized into two main groups: deterministic and stochastic. *Stochastic methods* are based on an element of random choice. Because of this, one has to sacrifice the possibility of an absolute guarantee of success within a finite amount of computation. *Deterministic methods*, such as the cutting plane, branch and bound methods, can find global optimal solutions, but not in polynomial time. Therefore, this type of methods can be used only for solving very small-sized problems. Indeed, global optimization problems with 200 variables are referred to as “medium scale”, problems with 1,000 variables as “large scale”, and the so-called “extra-large scale” is only around 4,000 variables [11]. In topology optimization, the variables could be easily 100 times more than this extra-large scale in global optimization. Therefore, to develop a unified deterministic theory for efficiently solving general global optimization problems is fundamentally important, not only in mathematical optimization, but also in general nonconvex analysis and mechanics.

2 Canonical Duality-Triality Theory

The canonical duality-triality theory comprises mainly three parts:

i) a *canonical dual transformation*, ii) a *complementary-dual principle*, and iii) a *triality theory*.

The canonical dual transformation is a versatile methodology which can be used to model complex systems within a unified framework and to formulate perfect dual problems without a duality gap. The complementary-dual principle presents a unified analytic solution form for general problems in continuous and discrete systems. The triality theory reveals an intrinsic duality pattern in multi-scale systems, which can be used to identify both global and local extrema, and to develop deterministic algorithms for ef-

fectively solving a wide class of nonconvex/nonsmooth/discrete optimization/variational problems.

2.1 General modeling and objectivity

A useful methodological theory should have solid foundations not only in physics, but also in mathematics, even in philosophy and aesthetics. The canonical duality theory was developed from Gao and Strang's original work for solving the following general nonconvex/nonsmooth variational problem [85]:

$$\min\{\Pi(\chi) = W(D\chi) - F(\chi) \mid \chi \in \mathcal{X}_c\}, \quad (3)$$

where $F(\chi)$ is the external energy, which must be linear on its domain \mathcal{X}_a ; the linear operator $D : \mathcal{X}_a \rightarrow \mathcal{W}_a$ assigns each configuration χ to an internal variable $\epsilon = D\chi$ and, correspondingly, $W : \mathcal{W}_a \rightarrow \mathbb{R}$ is called the internal (or stored) energy. The feasible set $\mathcal{X}_c = \{\chi \in \mathcal{X}_a \mid D\chi \in \mathcal{W}_a\}$ is the *kinetically admissible space*.

By Riesz representation theorem, the external energy can be written as $F(\chi) = \langle \chi, \bar{\chi}^* \rangle$, where $\bar{\chi}^* \in \mathcal{X}^*$ is a given input (or source). The bilinear form $\langle \chi, \chi^* \rangle : \mathcal{X} \times \mathcal{X}^* \rightarrow \mathbb{R}$ puts \mathcal{X} and \mathcal{X}^* in duality. Therefore, the variation (or Gâteaux derivative) of $F(\chi)$ leads to the *action-reaction duality*: $\bar{\chi}^* = \partial F(\chi)$. Dually, the internal energy must be an *objective function* on its domain \mathcal{W}_a such that the intrinsic physical behavior of the system can be described by the *constitutive duality*: $\sigma = \partial W(\epsilon)$.

Objectivity is a basic concept in mathematical modeling [19, 96, 118, 131], but is still subjected to seriously study in continuum physics [116, 126, 127]. The mathematical definition was given in Gao's book (Definition 6.1.2 [53]).

Definition 1 (Objectivity and Isotropy) Let \mathcal{R} be a proper orthogonal group, i.e. $\mathbf{R} \in \mathcal{R}$ if and only if $\mathbf{R}^T = \mathbf{R}^{-1}$, $\det \mathbf{R} = 1$. A set \mathcal{W}_a is said to be objective if

$$\mathbf{R}\epsilon \in \mathcal{W}_a \quad \forall \epsilon \in \mathcal{W}_a, \forall \mathbf{R} \in \mathcal{R}.$$

A real-valued function $W : \mathcal{W}_a \rightarrow \mathbb{R}$ is said to be objective if

$$W(\mathbf{R}\epsilon) = W(\epsilon) \quad \forall \epsilon \in \mathcal{W}_a, \forall \mathbf{R} \in \mathcal{R}. \quad (4)$$

A set \mathcal{W}_a is said to be isotropic if $\epsilon\mathbf{R} \in \mathcal{W}_a \quad \forall \epsilon \in \mathcal{W}_a, \forall \mathbf{R} \in \mathcal{R}$.

A real-valued function $W : \mathcal{W}_a \rightarrow \mathbb{R}$ is said to be isotropic if

$$W(\epsilon\mathbf{R}) = W(\epsilon) \quad \forall \epsilon \in \mathcal{W}_a, \forall \mathbf{R} \in \mathcal{R}. \quad (5)$$

Geometrically speaking, an objective function does not depend on the rotation, but only on certain measure of its variable. The isotropy means that the function $W(\epsilon)$ possesses a certain symmetry. In continuum physics, the right Cauchy-Green tensor³ $\mathbf{C}(\mathbf{F}) = \mathbf{F}^T \mathbf{F}$ is an objective strain measure, while the left Cauchy-Green tensor $\mathbf{c} = \mathbf{F} \mathbf{F}^T$ is an isotropic strain measure. In Euclidean space $\mathcal{W}_a \subset \mathbb{R}^n$, the simplest objective function is the ℓ_2 -norm $\|\epsilon\|$ in \mathbb{R}^n as we have $\|\mathbf{R}\epsilon\|^2 = \epsilon^T \mathbf{R}^T \mathbf{R} \epsilon = \|\epsilon\|^2 \quad \forall \mathbf{R} \in \mathcal{R}$. In

³Tensor is a geometrical object which is defined as a multi-dimensional array satisfying a transformation law (see [131]). A tensor must be independent of a particular choice of coordinate system (frame-indifference). But this terminology has been misused in optimization literature, where, any multi-dimensional array of data is called tensor (see [7]).

this case, the objectivity is equivalent to isotropy and, in Lagrangian mechanics, the kinetic energy is required to be isotropic [104].

Physically, an objective function doesn't depend on observers [127], which is essential for any real-world mathematical modelling. In continuum physics, objectivity implies that the equilibrium condition of angular momentum (symmetry of the Cauchy stress tensor $\sigma = \partial W(\epsilon)$, Section 6.1 [53]) holds. It is emphasized by P. Ciarlet that the objectivity is not an assumption, but an axiom [19]. Indeed, the objectivity is also known as the *axiom of material frame-invariance*, which lays a foundation for the canonical duality theory.

As an objective function, the internal energy $W(\epsilon)$ does not depend on each particular problem. Dually, the external energy $F(\chi)$ can be called the *subjective function*, which depends on each given problem, such as the inputs, boundary conditions and geometrical constraints in \mathcal{X}_a . Together, $\Pi(\chi) = W(D\chi) - F(\chi)$ is called the total potential energy and the minimal potential principle leads to the general optimization problem (3).

For dynamical problems, the linear operator $D = \{\partial_t, \partial_x\}$ and $W(D\chi) = T(\partial_t\chi) - V(\partial_x\chi)$, where $T(\mathbf{v})$ is the kinetic energy and $V(\mathbf{e})$ can be viewed as stored potential energy, then

$$\Pi(\chi) = T(\partial_t\chi) - V(\partial_x\chi) - F(\chi)$$

is the total action in dynamical systems.

The necessary condition $\delta\Pi(\chi) = 0$ for the solution of the minimization problem (3) leads to a general equilibrium equation:

$$A(\chi) = D^* \partial_\epsilon W(D\chi) = \bar{\chi}^*. \quad (6)$$

This abstract form of equilibrium equation covers extensive real-world applications ranging from traditional mathematical physics, modern economics, ecology, game theory, information technology, network optimization, operations research, and much more [53, 84, 145]. Particularly, if $W(\epsilon)$ is quadratic such that $\partial^2 W(\epsilon) = H$, then the operator $A : \mathcal{X}_c \rightarrow \mathcal{X}^*$ is linear and can be written in the triality form: $A = D^* H D$, which appears extensively in mathematical physics, optimization, and linear systems [53, 129, 145]. Clearly, any convex quadratic function $W(\epsilon)$ is objective due to the Cholesky decomposition $A = \Lambda^* \Lambda \succeq 0$.

Example 1 (Manufacturing/Production Systems) In management science, the configuration variable is a vector $\chi \in \mathbb{R}^n$, which could represent the products of a manufacture company. Its dual variable $\bar{\chi}^* \in \mathbb{R}^n$ can be considered as market price (or demands). Therefore, the external energy $F(\chi) = \langle \chi, \bar{\chi}^* \rangle = \chi^T \bar{\chi}^*$ in this example is the total income of the company. The products are produced by workers $\epsilon \in \mathbb{R}^m$. Due to the cooperation, we have $\epsilon = D\chi$ and $D \in \mathbb{R}^{m \times n}$ is a matrix. Workers are paid by salary $\sigma = \partial W(\epsilon)$, therefore, the internal energy $W(\epsilon)$ in this example is the cost, which should be an objective function. Thus, $\Pi(\chi) = W(D\chi) - F(\chi)$ is the *total cost or target* and the minimization problem $\min \Pi(\chi)$ leads to the equilibrium equation

$$D^T \partial_\epsilon W(D\chi) = \bar{\chi}^*,$$

which is an algebraic equation in \mathbb{R}^n . The weak form of this equilibrium equation is $\langle \chi, D^T \sigma \rangle = \langle D\chi, \sigma \rangle = \langle \chi, \bar{\chi}^* \rangle$, which is the well-known *D'Alembert's principle* or the *principle of virtual work* in Lagrangian mechanics. The cost function $W(\epsilon)$ could be

convex for a very small company, but usually nonconvex for big companies to allow some people having the same salaries.

Example 2 (Lagrange Mechanics) In analytical mechanics, the configuration $\chi \in \mathcal{X}_a \subset \mathcal{C}^1[I; \mathbb{R}^n]$ is a continuous vector-valued function of time $t \in I \subset \mathbb{R}$. Its components $\{\chi_i\}$ ($i = 1, \dots, n$) are known as the *Lagrangian coordinates*⁴. Its dual variable $\bar{\chi}^*$ is the action vector function in \mathbb{R}^n , say $\mathbf{f}(t)$. The external energy $F(\chi) = \langle \chi, \bar{\chi}^* \rangle = \int_I \chi(t) \cdot \mathbf{f}(t) dt$. While the internal energy $W(D\chi)$ is the so-called action:

$$W(D\chi) = \int_I L(\chi, \dot{\chi}) dt, \quad L = T(\dot{\chi}) - V(\chi)$$

where T is the kinetic energy density, V is the potential density, and $L = T - V$ is the standard *Lagrangian density*. In this case, the linear operator $D\chi = \{\partial_t, 1\}\chi = \{\dot{\chi}, \chi\}$ is a vector-valued mapping. The kinetic energy T must be an objective function of the velocity $\mathbf{v}_k = \dot{\chi}_k(\chi)$ (or isotropic since \mathbf{v}_k is a vector) of each particle $\mathbf{x}_k = \mathbf{x}_k(\chi) \in \mathbb{R}^3 \quad \forall k \in I_m$, while the potential density V depends on each problem. Together, $\Pi(\chi) = W(D\chi) - F(\chi)$ is called *total action*. Its stationary condition leads to the *Euler-Lagrange equation*:

$$D^* \partial W(D\chi) = -\partial_t \frac{\partial T(\dot{\chi})}{\partial \dot{\chi}} - \nabla V(\chi) = \mathbf{f}. \quad (7)$$

For Newton mechanics, $T(\mathbf{v}) = \frac{1}{2} \sum_{k \in I_m} m_k \|\mathbf{v}_k\|^2$ is quadratic, where $\|\mathbf{v}_k\|$ represents the Euclidean norm (speed) of the k -th particle in \mathbb{R}^3 . For Einstein's special relativity theory, $T(\mathbf{v}) = -m_0 c \sqrt{c^2 - \|\mathbf{v}\|^2}$ is convex (see Chapter 2.1.2, [53]), where $m_0 > 0$ is the mass of a particle at rest, c is the speed of light. Therefore, the total action $\Pi(\chi)$ is convex only if $V(\chi)$ is linear. In this case, the solution of the Euler-Lagrange equation (7) minimizes the total action. The total action is nonconvex as long as the potential density $V(\chi)$ is nonlinear. In this case, the system may have periodic solution if $V(\chi)$ is convex and the well-known *least action principle is indeed a misnomer* (see Chapter 2, [53]). The system may have chaotic solution if the potential density $V(\chi)$ is nonconvex [54, 61]. Unfortunately, these important facts are not well-realized in both classical mechanics and modern nonlinear dynamical systems. The recent review article [71] presents a unified understanding bifurcation, chaos, and NP-hard problems in complex systems.

In nonlinear analysis, the linear operator D is a partial differential operator, say $D = \{\partial_t, \partial_x\}$, and the abstract equilibrium equation (6) is a nonlinear partial differential equation. For convex $W(\epsilon)$, the solution of this equilibrium equation is also a solution to the minimization problem (3). However, for nonconvex $W(\epsilon)$, the solution of (6) is only a stationary point of $\Pi(\chi)$. In order to study stability and regularity of the local solutions in nonconvex problems, many generalized definitions, such as quasi-, poly- and rank-one convexities have been introduced and subjected to extensively study for more than fifty years [8]. But all these generalized convexities provide only local extremality conditions, which lead to many “outstanding open problems” in nonlinear analysis [8]. However, by the canonical duality-triality theory, we can have clear understandings on these challenges.

⁴It is an unfortunate truth that many people don't know the relation between the Lagrangian space \mathbb{R}^n they work in and the Minkowski (physical) space $\mathbb{R}^3 \times \mathbb{R}$ they live in.

2.2 Canonical transformation and classification of nonlinearities

According to the canonical duality, the linear measure $\epsilon = D\chi$ can't be used directly for studying constitutive law due to the objectivity. Also, the linear operator can't change the nonconvexity of $W(D\chi)$. Indeed, it is well-known that the deformation gradient $\mathbf{F} = \nabla\chi$ is not considered as a strain measure in nonlinear elasticity. The most commonly used strain measure is the right Cauchy-Green strain tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$, which is, clearly, an objective function since $\mathbf{C}(\mathbf{F}) = \mathbf{C}(\mathbf{QF})$. According to P. Ciarlet (Theorem 4.2-1, [18]), the stored energy $W(\mathbf{F})$ of a hyperelastic material is objective if and only if there exists a function \tilde{W} such that $W(\mathbf{F}) = \tilde{W}(\mathbf{C})$. Based on this fact in continuum physics, the canonical transformation is naturally introduced.

Definition 2 (Canonical Function and Canonical Transformation)

A real-valued function $\Phi : \mathcal{E}_a \rightarrow \mathbb{R}$ is called canonical if the duality mapping $\partial\Phi : \mathcal{E}_a \rightarrow \mathcal{E}_a^*$ is one-to-one and onto.

For a given nonconvex function $W : \mathcal{W}_a \rightarrow \mathbb{R}$, if there exists a geometrically admissible mapping $\Lambda : \mathcal{W}_a \rightarrow \mathcal{E}_a$ and a canonical function $\Phi : \mathcal{E}_a \rightarrow \mathbb{R}$ such that

$$W(\epsilon) = \Phi(\Lambda(\epsilon)), \quad (8)$$

then, the transformation (8) is called the canonical transformation and $\xi = \Lambda(\epsilon)$ is called the canonical measure.

By this definition, the one-to-one duality relation $\xi^* = \partial\Phi(\xi) : \mathcal{E}_a \rightarrow \mathcal{E}_a^*$ implies that the canonical function $\Phi(\xi)$ is differentiable and its conjugate function $\Phi^* : \mathcal{E}_a^* \rightarrow \mathbb{R}$ can be uniquely defined by the Legendre transformation [53]

$$\Phi^*(\xi^*) = \{\langle \xi; \xi^* \rangle - \Phi(\xi) \mid \xi^* = \partial\Phi(\xi)\}, \quad (9)$$

where $\langle \xi; \xi^* \rangle$ represents the bilinear form on \mathcal{E} and its dual space \mathcal{E}^* . In this case, $\Phi : \mathcal{E}_a \rightarrow \mathbb{R}$ is a canonical function if and only if the following canonical duality relations hold on $\mathcal{E}_a \times \mathcal{E}_a^*$:

$$\xi^* = \partial\Phi(\xi) \Leftrightarrow \xi = \partial\Phi^*(\xi^*) \Leftrightarrow \Phi(\xi) + \Phi^*(\xi^*) = \langle \xi; \xi^* \rangle. \quad (10)$$

A canonical function $\Phi(\xi)$ can also be nonsmooth but should be convex such that its conjugate can be well-defined by Fenchel transformation

$$\Phi^\sharp(\xi^*) = \sup\{\langle \xi; \xi^* \rangle - \Phi(\xi) \mid \xi \in \mathcal{E}_a\}. \quad (11)$$

In this case, $\partial\Phi(\xi) \subset \mathcal{E}_a^*$ is understood as the sub-differential and the canonical duality relations (10) should be written in the generalized form

$$\xi^* \in \partial\Phi(\xi) \Leftrightarrow \xi \in \partial\Phi^\sharp(\xi^*) \Leftrightarrow \Phi(\xi) + \Phi^\sharp(\xi^*) = \langle \xi; \xi^* \rangle. \quad (12)$$

This generalized canonical duality plays an important role in unified understanding Lagrangian duality and KKT theory for constrained optimization problems (see [82, 107] and Section 5.4).

In analysis, nonlinear PDEs are classified as semilinear, quasi-linear, and fully nonlinear three categories based on the degree of the nonlinearity [29]. A *semilinear PDE* is a differential equation that is nonlinear in the unknown function but linear in all

its partial derivatives. A *quasi-linear PDE* is one that is nonlinear in (at least) one of the lower order derivatives but linear in the highest order derivative(s) of the unknown function. *Fully nonlinear PDEs* are referred to as the class of nonlinear PDEs which are nonlinear in the highest order derivatives of the unknown function. However, this classification is not essential as we know that the main difficulty is nonconvexity, instead of nonlinearity since these nonlinear PDEs could be related to certain convex variational problems, which can be solved easily by numerical methods.

The concepts of geometrical and physical nonlinearities are well-known in continuum physics, but not in abstract analysis and optimization. This leads to many confusions. Based on the canonical transformation, we can have the following classification.

Definition 3 (Geometrical, Physical and Complete Nonlinearities)

The general problem (3) is called geometrically nonlinear (resp. linear) if the geometrical operator $\Lambda(\epsilon)$ is nonlinear (resp. linear);

The problem (3) is called physically nonlinear (resp. linear) if the constitutive relation $\xi^ = \partial\Phi(\xi)$ is nonlinear (resp. linear);*

The general problem (3) is called completely nonlinear if it is both geometrically and physically nonlinear.

According to this clarification, the minimization problem (3) is geometrically linear as long as the stored energy $W(\epsilon)$ is convex. In this case, $\Lambda(D\chi) = D\chi$ and $\Phi(\Lambda(\epsilon)) = W(\epsilon)$. Thus, a physically nonlinear but geometrically linear problem could be equivalent to a fully nonlinear PDE, which can be solved easily by well-developed convex optimization techniques. Therefore, the main difficulty in complex systems is the geometrical nonlinearity. This is the reason why only this nonlinearity was emphasized in the title of Gao-Strang's paper [85]. The complete nonlinearity is also called fully nonlinearity in engineering mechanics. Hope this new classification will clear out this confusion. By the canonical transformation, the completely nonlinear minimization problem (3) can be equivalently written in the following canonical form

$$(\mathcal{P}) : \min\{\Pi(\chi) = \Phi(\Lambda(D\chi)) - F(\chi) \mid \chi \in \mathcal{X}_c\}. \quad (13)$$

In order to solving this nonconvex problem, we need to find its canonical dual form.

2.3 Complementary-dual principle

For geometrically linear problems, the stored energy $W(\epsilon)$ is convex and the complementary energy $W^*(\sigma)$ can be uniquely defined on \mathcal{W}_a^* by Legendre transformation. Therefore, by using equality $W(\epsilon) = \langle \epsilon; \sigma \rangle - W^*(\sigma)$, the total potential $\Pi(\chi)$ can be equivalently written in the classical Lagrangian form $L : \mathcal{X}_a \times \mathcal{W}_a^* \rightarrow \mathbb{R}$

$$L(\chi, \sigma) = \langle D\chi; \sigma \rangle - W^*(\sigma) - F(\chi) = \langle \chi, D^*\sigma - \bar{\chi}^* \rangle - W^*(\sigma), \quad (14)$$

where, χ can be viewed as a Lagrange multiplier for the equilibrium equation $D^*\sigma = \bar{\chi}^*$. In linear elasticity, $L(\chi, \sigma)$ is the well-known Hellinger-Reissner complementary energy. Let $\mathcal{S}_c = \{\sigma \in \mathcal{W}_a^* \mid D^*\sigma = \bar{\chi}^*\}$ be the so-called *statically admissible space*. Then the Lagrangian dual of the general problem (3) is given by

$$\max\{\Pi^*(\sigma) = -W^*(\sigma) \mid \sigma \in \mathcal{S}_c\}, \quad (15)$$

and the following Lagrangian min-max duality is well-known:

$$\min_{\chi \in \mathcal{X}_c} \Pi(\chi) = \min_{\chi \in \mathcal{X}_a} \max_{\sigma \in \mathcal{W}_a^*} L(\chi, \sigma) = \max_{\sigma \in \mathcal{W}_a^*} \min_{\chi \in \mathcal{X}_a} L(\chi, \sigma) = \max_{\sigma \in \mathcal{S}_c} \Pi^*(\sigma). \quad (16)$$

In continuum mechanics, this one-to-one duality is called *complementary-dual variational principle* [129]. In finite elasticity, the Lagrangian dual is also known as the *Levison-Zubov principle*. However, this principle holds only for convex problems. If the stored energy $W(\epsilon)$ is nonconvex, its complementary energy can't be determined uniquely by the Legendre transformation. Although its Fenchel conjugate $W^\sharp : \mathcal{W}_a^* \rightarrow \mathbb{R} \cup \{+\infty\}$ can be uniquely defined, the Fenchel-Moreau dual problem

$$\max\{\Pi^\sharp(\sigma) = -W^\sharp(\sigma) \mid \sigma \in \mathcal{S}_c\} \quad (17)$$

is not considered as a complementary-dual problem due to Fenchel-Young inequality:

$$\min\{\Pi(\chi) \mid \chi \in \mathcal{X}_c\} \geq \max\{\Pi^\sharp(\sigma) \mid \sigma \in \mathcal{S}_c\}, \quad (18)$$

and $\theta = \min \Pi(\chi) - \max \Pi^\sharp(\sigma) \neq 0$ is the so-called *duality gap*. This duality gap is intrinsic to all type of Lagrangian duality problems since the nonconvexity of $W(D\chi)$ can't be changed by any linear operator. It turns out that the existence of a pure stress based complementary-dual principle has been a well-known debt in finite elasticity for more than forty years [114].

Remark 1 (Lagrange Multiplier Law) Strictly speaking, the Lagrange multiplier method can be used mainly for equilibrium constraint in \mathcal{S}_c and the Lagrange multiplier must be the solution to the primal problem (see Section 1.5.2 [53]). The equilibrium equation $D^*\sigma = \bar{\chi}^*$ must be an invariant under certain coordinates transformation, say the law of angular momentum conservation, which is guaranteed by the objectivity of the stored energy $W(D\chi)$ in continuum mechanics (see Definition 6.1.2, [53]), or by the isotropy of the kinetic energy $T(\dot{\chi})$ in Lagrangian mechanics [104]. Specifically, the equilibrium equation for Newton's mechanics is an invariant under the Calilean transformation; while for Einstein's special relativity theory, the equilibrium equation $D^*\sigma = \bar{\chi}^*$ is an invariant under the Lorentz transformation. For linear equilibrium equation, the quadratic $W(\epsilon)$ is naturally an objective function for convex systems. Unfortunately, since the concept of the objectivity is misused in mathematical optimization, the Lagrange multiplier method has been mistakenly used for solving general nonconvex problems, which produces many different duality gaps.

In order to recover the duality gap in nonconvex problems, we use the canonical transformation $W(D\chi) = \Phi(\Lambda(D\chi))$ such that the nonconvex total potential $\Pi(\chi)$ can be reformulated as the total complementary energy $\Xi : \mathcal{X}_a \times \mathcal{E}_a^* \rightarrow \mathbb{R}$

$$\Xi(\chi, \xi^*) = \langle \Lambda(D\chi); \xi^* \rangle - \Phi^*(\xi^*) - F(\chi), \quad (19)$$

which was first introduced by Gao and Strang in 1989 [85]. The stationary condition $\delta\Xi(\chi, \xi^*) = 0$ leads to the following canonical equations:

$$\Lambda(D\chi) = \partial\Phi^*(\xi^*), \quad (20)$$

$$D^*\Lambda_t(D\chi)\xi^* = \partial F(\chi), \quad (21)$$

where $\Lambda_t(\epsilon) = \partial\Lambda(\epsilon)$ is a generalized Gâteaux derivative of $\Lambda(\epsilon)$. By the canonical duality, (20) is equivalent to $\xi^* = \partial_{\xi}\Phi(\Lambda(D\chi))$. Therefore, the canonical equilibrium equation (21) is the general equilibrium equation (6).

By using the Gao-Strang complementary function, the canonical dual of $\Pi(\chi)$ can be obtained as

$$\Pi^d(\xi^*) = \text{sta}\{\Xi(\chi, \xi^*) \mid \chi \in \mathcal{X}_a\} = F^\Lambda(\xi^*) - \Phi^*(\xi^*), \quad (22)$$

where $F^\Lambda(\xi^*)$ is the Λ -transformation defined by [55]

$$F^\Lambda(\xi^*) = \text{sta}\{\langle \Lambda(D\chi); \xi^* \rangle - F(\chi) \mid \chi \in \mathcal{X}_a\}. \quad (23)$$

Clearly, the stationary condition in this Λ -transformation is the canonical equilibrium equation (21). Let $\mathcal{S}_c \subset \mathcal{E}_a^*$ be a feasible set, on which $F^\Lambda(\xi^*)$ is well-defined. Then we have the following result.

Theorem 1 (Complementary-Dual Principle [49, 51]) *If $(\bar{\chi}, \bar{\xi}^*) \in \mathcal{X}_a \times \mathcal{E}_a^*$ is a stationary point of $\Xi(\chi, \xi^*)$, then $\bar{\chi}$ is a stationary point of $\Pi(\chi)$ on \mathcal{X}_c , while $\bar{\xi}^*$ is a stationary point of $\Pi^d(\xi^*)$ on \mathcal{S}_c , and*

$$\Pi(\bar{\chi}) = \Xi(\bar{\chi}, \bar{\xi}^*) = \Pi^d(\bar{\xi}^*). \quad (24)$$

This theorem shows that there is no duality gap between $\Pi(\chi)$ and $\Pi^d(\xi^*)$. In many real-world applications, the geometrical operator $\Lambda(\epsilon)$ is usually quadratic such that the total complementary function $\Xi(\chi, \xi^*)$ can be written as

$$\Xi(\chi, \xi^*) = \frac{1}{2}\langle \chi, \mathbf{G}(\xi^*)\chi \rangle - \Phi^*(\xi^*) - \langle \chi, \mathbf{F}(\xi^*) \rangle \quad (25)$$

where $\mathbf{G}(\xi^*) = \nabla_{\chi}^2 \Xi(\chi, \xi^*)$ and $\mathbf{F}(\xi^*)$ depends on the linear terms in $\Lambda(D\chi)$ and the input $\bar{\chi}^*$. The first term in $\Xi(\chi, \xi^*)$

$$G_{ap}(\chi, \xi^*) = \frac{1}{2}\langle \chi, \mathbf{G}(\xi^*)\chi \rangle \quad (26)$$

is the so-called *complementary gap function* introduced by Gao and Strang in [85]. In this case, the canonical equilibrium equation $\nabla_{\chi} \Xi(\chi, \xi^*) = \mathbf{G}(\xi^*)\chi - \mathbf{F}(\xi^*) = 0$ is linear in χ and the canonical dual Π^d can be explicitly formulated as

$$\Pi^d(\xi^*) = -G_{ap}^*(\xi^*) - \Phi^*(\xi^*), \quad (27)$$

where $G_{ap}^*(\xi^*) = \frac{1}{2}\langle \mathbf{G}^{-1}(\xi^*)\mathbf{F}(\xi^*), \mathbf{F}(\xi^*) \rangle$ is called *pure complementary gap function*. Comparing this canonical dual with the Lagrangian dual $\Pi^*(\sigma) = -W^*(\sigma)$ in (15) we can find that in addition to replace W^* by the canonical dual Φ^* , the first term in Π^d is identical to the Gao-Strang complementary gap function, which recovers the duality gap in Lagrangian duality theory and plays an important role in triality theory.

Theorem 2 (Analytical Solution Form) *If $\bar{\xi}^* \in \mathcal{S}_c$ is a stationary point of $\Pi^d(\xi^*)$, then*

$$\bar{\chi} = \mathbf{G}^{-1}(\bar{\xi}^*)\mathbf{F}(\bar{\xi}^*) \quad (28)$$

is a stationary point of $\Pi(\chi)$ on \mathcal{X}_c and $\Pi(\bar{\chi}) = \Pi^d(\bar{\xi}^)$.*

This theorem shows that the primal solution is analytically depends on its canonical dual solution. Clearly, the canonical dual of a nonconvex primal problem is also non-convex and may have multiple stationary points. By the canonical duality, each of these stationary solutions is corresponding to a primal solution via (28). Their extremality is governed by Gao and Strang's complementary gap function.

2.4 Triality theory

In order to identify extremality of these stationary solutions, we need to assume that the canonical function $\Phi : \mathcal{E}_a \rightarrow \mathbb{R}$ is convex and let

$$\mathcal{S}_c^+ = \{\xi^* \in \mathcal{S}_c \mid \mathbf{G}(\xi^*) \succ 0\}, \quad \mathcal{S}_c^- = \{\xi^* \in \mathcal{S}_c \mid \mathbf{G}(\xi^*) \prec 0\}. \quad (29)$$

Clearly, for any given $\chi \in \mathcal{X}_a$ and $\chi \neq 0$, we have

$$G_{ap}(\chi, \xi^*) > 0 \Leftrightarrow \xi^* \in \mathcal{S}_c^+, \quad G_{ap}(\chi, \xi^*) < 0 \Leftrightarrow \xi^* \in \mathcal{S}_c^-.$$

Theorem 3 (Triality Theorem) Suppose $\bar{\xi}^*$ is a stationary point of $\Pi^d(\xi^*)$ and $\bar{\chi} = \mathbf{G}^{-1}(\bar{\xi}^*)\bar{\xi}^*$. If $\bar{\xi}^* \in \mathcal{S}_c^+$, we have

$$\Pi(\bar{\chi}) = \min_{\chi \in \mathcal{X}_c} \Pi(\chi) \Leftrightarrow \max_{\xi^* \in \mathcal{S}_c^+} \Pi^d(\xi^*) = \Pi^d(\bar{\xi}^*); \quad (30)$$

If $\bar{\xi}^* \in \mathcal{S}_c^-$, then on a neighborhood⁵ $\mathcal{X}_o \times \mathcal{S}_o \subset \mathcal{X}_c \times \mathcal{S}_c^-$ of $(\bar{\chi}, \bar{\xi}^*)$, we have either

$$\Pi(\bar{\chi}) = \max_{\chi \in \mathcal{X}_o} \Pi(\chi) \Leftrightarrow \max_{\xi^* \in \mathcal{S}_o} \Pi^d(\xi^*) = \Pi^d(\bar{\xi}^*), \quad (31)$$

or (only if $\dim \bar{\chi} = \dim \bar{\xi}^*$)

$$\Pi(\bar{\chi}) = \min_{\chi \in \mathcal{X}_o} \Pi(\chi) \Leftrightarrow \min_{\xi^* \in \mathcal{S}_o} \Pi^d(\xi^*) = \Pi^d(\bar{\xi}^*). \quad (32)$$

The first statement (30) is called *canonical min-max duality*. Its weak form was discovered by Gao and Strang in 1989 [85]. This duality can be used to identify global minimizer of the nonconvex problem (3). According this statement, the nonconvex problem (3) is equivalent to the following canonical dual problem, denoted by (\mathcal{P}^d) :

$$(\mathcal{P}^d) : \max\{\Pi^d(\xi^*) \mid \xi^* \in \mathcal{S}_c^+\}. \quad (33)$$

This is a concave maximization problem which can be solved easily by well-developed convex analysis and optimization techniques. The second statement (31) is the *canonical double-max duality* and (32) is the *canonical double-min duality*. These two statements can be used to identify the biggest local maximizer and local minimizer of the primal problem, respectively.

The triality theory was first discovered by Gao 1996 in post-buckling analysis of a large deformed beam [46]. The generalization to global optimization was made in 2000 [55]. It was realized in 2003 that the double-min duality (32) holds under certain additional condition [61, 62]. Recently, it is proved that this additional condition is simply $\dim \bar{\chi} = \dim \bar{\xi}^*$ to have the strong canonical double-min duality (32), otherwise, this double-min duality holds weakly in subspaces of $\mathcal{X}_o \times \mathcal{S}_o$ [88, 89, 119, 120].

Example 3 To explain the theory, let us consider a very simple nonconvex optimization in \mathbb{R}^n :

$$\min \left\{ \Pi(\mathbf{x}) = \frac{1}{2}\alpha \left(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda \right)^2 - \mathbf{x}^T \mathbf{f} \quad \forall \mathbf{x} \in \mathbb{R}^n \right\}, \quad (34)$$

⁵The neighborhood \mathcal{X}_o of $\bar{\chi}$ means that on which, $\bar{\chi}$ is the only stationary point.

where $\alpha, \lambda > 0$ are given parameters. The criticality condition $\nabla P(\mathbf{x}) = 0$ leads to a nonlinear algebraic equation system in \mathbb{R}^n

$$\alpha\left(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda\right)\mathbf{x} = \mathbf{f}. \quad (35)$$

Clearly, to solve this nonlinear algebraic equation directly is difficult. Also traditional convex optimization theory can not be used to identify global minimizer. However, by the canonical dual transformation, this problem can be solved completely and easily. To do so, we let $\xi = \Lambda(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}\|^2 \in \mathbb{R}$, which is an objective measure. Then, the nonconvex function $W(\mathbf{x}) = \frac{1}{2}\alpha(\frac{1}{2}\|\mathbf{x}\|^2 - \lambda)^2$ can be written in canonical form $\Phi(\xi) = \frac{1}{2}\alpha(\xi - \lambda)^2$. Its Legendre conjugate is given by $\Phi^*(\varsigma) = \frac{1}{2}\alpha^{-1}\varsigma^2 + \lambda\varsigma$, which is strictly convex. Thus, the total complementary function for this nonconvex optimization problem is

$$\Xi(\mathbf{x}, \varsigma) = \frac{1}{2}\|\mathbf{x}\|^2\varsigma - \frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma - \mathbf{x}^T\mathbf{f}. \quad (36)$$

For a fixed $\varsigma \in \mathbb{R}$, the criticality condition $\nabla_{\mathbf{x}}\Xi(\mathbf{x}) = 0$ leads to

$$\varsigma\mathbf{x} - \mathbf{f} = 0. \quad (37)$$

For each $\varsigma \neq 0$, the equation (37) gives $\mathbf{x} = \mathbf{f}/\varsigma$ in vector form. Substituting this into the total complementary function Ξ , the canonical dual function can be easily obtained as

$$\Pi^d(\varsigma) = \{\Xi(\mathbf{x}, \varsigma) | \nabla_{\mathbf{x}}\Xi(\mathbf{x}, \varsigma) = 0\} = -\frac{\mathbf{f}^T\mathbf{f}}{2\varsigma} - \frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma, \quad \forall \varsigma \neq 0. \quad (38)$$

The critical point of this canonical function is obtained by solving the following dual algebraic equation

$$2(\alpha^{-1}\varsigma + \lambda)\varsigma^2 = \mathbf{f}^T\mathbf{f}. \quad (39)$$

For any given parameters α, λ and the vector $\mathbf{f} \in \mathbb{R}^n$, this cubic algebraic equation has at most three real roots satisfying $\varsigma_1 \geq 0 \geq \varsigma_2 \geq \varsigma_3$, and each of these roots leads to a critical point of the nonconvex function $P(\mathbf{x})$, i.e., $\mathbf{x}_i = \mathbf{f}/\varsigma_i$, $i = 1, 2, 3$. By the fact that $\varsigma_1 \in \mathcal{S}_c^+ = \{\varsigma \in \mathbb{R} \mid \varsigma > 0\}$, $\varsigma_{2,3} \in \mathcal{S}_c^- = \{\varsigma \in \mathbb{R} \mid \varsigma < 0\}$, then Theorem 3 tells us that \mathbf{x}_1 is a global minimizer of $\Pi(\mathbf{x})$, \mathbf{x}_3 is a local maximizer of $\Pi(\mathbf{x})$, while \mathbf{x}_2 is a local minimizer if $n = 1$ (see Fig. 1). If we choose $n = 1$, $\alpha = 1$, $\lambda = 2$, and $f = \frac{1}{2}$, the primal function and canonical dual function are shown in Fig. 1 (a), where, $x_1 = 2.11491$ is global minimizer of $\Pi(\mathbf{x})$, $\varsigma_1 = 0.236417$ is global maximizer of $\Pi^d(\varsigma)$, and $\Pi(x_1) = -1.02951 = \Pi^d(\varsigma_1)$ (see the two black dots). Also it is easy to verify that x_2 is a local minimizer, while x_3 is a local maximizer.

If we let $\mathbf{f} = 0$, the graph of $\Pi(\mathbf{x})$ is symmetric (i.e. the so-called double-well potential or the Mexican hat for $n = 2$ [61]) with infinite number of global minimizers satisfying $\|\mathbf{x}\|^2 = 2\lambda$. In this case, the canonical dual $\Pi^d(\varsigma) = -\frac{1}{2}\alpha^{-1}\varsigma^2 - \lambda\varsigma$ is strictly concave with only one critical point (local maximizer) $\varsigma_3 = -\alpha\lambda \in \mathcal{S}_c^-$ (for $\alpha, \lambda > 0$). The corresponding solution $\mathbf{x}_3 = \mathbf{f}/\varsigma_3 = 0$ is a local maximizer. By the canonical dual equation (39) we have $\varsigma_1 = \varsigma_2 = 0$ located on the boundary of \mathcal{S}_c^+ , which corresponding to the two global minimizers $x_{1,2} = \pm\sqrt{2\lambda}$ for $n = 1$, see Fig. 1 (b). This is similar to the post-buckling of large deformed beam. Due to symmetry ($f = 0$), the nonconvex function $\Pi(\mathbf{x})$ has two possible buckled solutions $\mathbf{x}_{1,2} = (\pm\sqrt{2\lambda}, 0)$ with the axial load

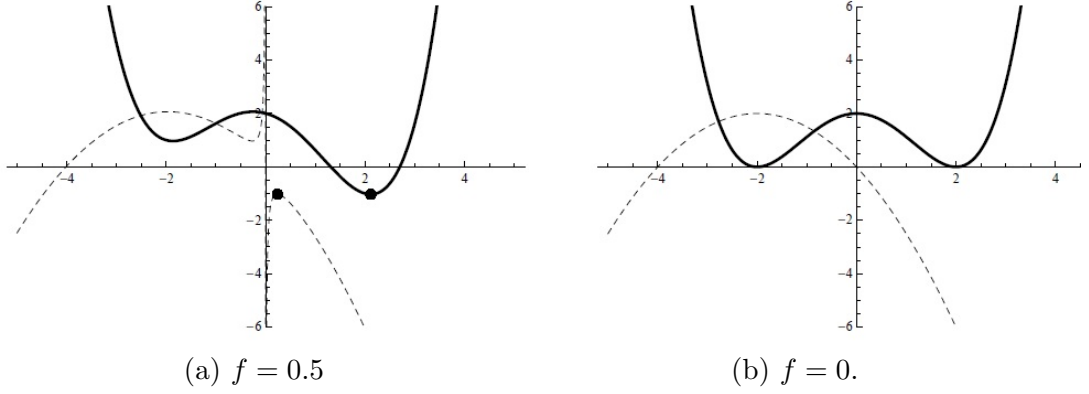


Figure 1: Graphs of $\Pi(\mathbf{x})$ (solid) and $\Pi^d(\zeta)$ (dashed)

$\lambda = \frac{1}{2}(b^2 - a^2)$. While the local maximizer $\mathbf{x}_3 = \{0, 0\}$ is corresponding to the unbuckled state.

This simple example shows a fundamental issue in global optimization, i.e., the optimal solutions of a nonconvex problem depends sensitively on the linear term (input or perturbation) \mathbf{f} . Geometrically speaking, the objective function $W(D\mathbf{x})$ in $\Pi(\mathbf{x})$ possesses certain symmetry. If there is no linear term (subjective function) in $\Pi(\mathbf{x})$, the nonconvex problem usually has more than one global minimizer due to the symmetry. Traditional direct approaches and the popular SDP method are usually failed to deal with this situation. By the canonical duality theory, we understand that in this case the canonical dual function $\Pi^d(\zeta)$ has no critical point in \mathcal{S}_c^+ . Therefore, the input \mathbf{f} breaks the symmetry so that $\Pi^d(\zeta)$ has a unique stationary point in \mathcal{S}_c^+ which can be obtained easily. This idea was originally from Gao's work (1996) on post-buckling analysis of large deformed beam [43], where the triality theorem was first proposed [46]. The potential energy of this beam model is a double-well function, similar to this example, without lateral force or imperfection, the beam could have two buckling states (corresponding to two minimizers) and one un-buckled state (local maximizer). Later on (2008) in the Gao and Ogden work on analytical solutions in phase transformation [75], they further discovered that the nonconvex system has no phase transition unless the force distribution $f(x)$ vanished at certain points. They also discovered that if force field $f(x)$ changes dramatically, all the Newton type direct approaches failed even to find any local minimizer. The linear perturbation method has been used successfully for solving global optimization problems [16, 120, 142, 155].

3 Applications for modeling of complex systems

By the fact that the canonical duality is a fundamental law governing natural phenomena and the objectivity is a basic condition for mathematical models, the canonical duality-triality theory can be used for modeling real-world problems within a unified framework.

3.1 Mixed integer nonlinear programming

The most general and challenging problem in global optimization could be the mixed integer nonlinear program (MINP), which is a minimization problem generally formulated

as (see [94])

$$\min\{f(\mathbf{x}, \mathbf{y}) \mid g_i(\mathbf{x}, \mathbf{y}) \leq 0 \ \forall i \in I_m, \ \mathbf{x} \in \mathbb{R}^n, \ \mathbf{y} \in \mathbb{Z}^p\} \quad (40)$$

where \mathbb{Z}^p is an integer set, the “objective function” $f(\mathbf{x}, \mathbf{y})$ and constraints $g_i(\mathbf{x}, \mathbf{y})$ for $i \in I_m$ are arbitrary functions [12]. Certainly, this artificial model is virtually applicable to any problem in operations research, but it is impossible to develop a general theory and powerful algorithm without detailed information given on these functions. As we know that the objectivity is a fundamental concept in mathematical modeling. Unfortunately, this concept has been mistakenly used with other functions, such as target, cost, energy, and utility functions, etc⁶.

Based on the Gao-Strang model (3), we let $\boldsymbol{\chi} = (\mathbf{x}, \mathbf{y})$, $\mathbf{D}\boldsymbol{\chi} = (\mathbf{D}_x\mathbf{x}, \mathbf{D}_y\mathbf{y})$, and $\bar{\boldsymbol{\chi}}^* = (\mathbf{b}, \mathbf{t})$. Then the general MINP problem (40) can be remodeled in the following form

$$\min\{\Pi(\mathbf{x}, \mathbf{y}) = W(\mathbf{D}_x\mathbf{x}, \mathbf{D}_y\mathbf{y}) - \mathbf{x}^T\mathbf{b} - \mathbf{y}^T\mathbf{t} \mid (\mathbf{x}, \mathbf{y}) \in \mathcal{X}_c \times \mathcal{Y}_c, \ \mathbf{x} \in \mathbb{Z}^p\}, \quad (41)$$

where the feasible sets are, correspondingly,

$$\mathcal{X}_c = \{\mathbf{x} \in \mathcal{X}_a \subset \mathbb{R}^n \mid \mathbf{D}_x\mathbf{x} \in \mathcal{U}_a\}, \ \mathcal{Y}_c = \{\mathbf{y} \in \mathcal{Y}_a \subset \mathbb{R}^p \mid \mathbf{D}_y\mathbf{y} \in \mathcal{V}_a\}.$$

In $\mathcal{X}_a, \mathcal{Y}_a$, certain linear constraints are given, while in $\mathcal{U}_a, \mathcal{V}_a$, general nonlinear (constitutive) constraints are prescribed such that the nonconvex (objective) function $W : \mathcal{U}_a \times \mathcal{V}_a \rightarrow \mathbb{R}$ can be written in the canonical form $W(\mathbf{D}\boldsymbol{\chi}) = \Phi_\chi(\boldsymbol{\Lambda}(\boldsymbol{\chi}))$ for certain geometrical operator $\boldsymbol{\Lambda}(\boldsymbol{\chi})$. By the fact that any integer set \mathbb{Z}^p is equivalent to a Boolean set [141, 154], we simply let $\mathbb{Z}^p = \{0, 1\}^p$. This constitutive constraint can be relaxed by the canonical transformation [68, 80]

$$\boldsymbol{\epsilon} = \boldsymbol{\Lambda}_x(\mathbf{x}) = \mathbf{x} \circ (\mathbf{x} - \mathbf{1}) = \{x_i^2 - x_i\}^p, \quad (42)$$

and the canonical function $\Phi_x(\boldsymbol{\epsilon}) = \{0 \text{ if } \boldsymbol{\epsilon} = \mathbf{0}, \infty \text{ otherwise}\}$. Therefore, the canonical form for the MINP problem is

$$\min\{\Pi(\mathbf{x}, \mathbf{y}) = \Phi_\chi(\boldsymbol{\Lambda}(\mathbf{x}, \mathbf{y})) + \Phi_x(\boldsymbol{\Lambda}_x(\mathbf{x})) - \mathbf{x}^T\mathbf{b} - \mathbf{y}^T\mathbf{t} \mid (\mathbf{x}, \mathbf{y}) \in \mathcal{X}_c \times \mathcal{Y}_c\}. \quad (43)$$

This canonical form covers many real-world applications, including the so-called fixed cost problem [83]. By the fact that the canonical function $\Phi_x(\boldsymbol{\epsilon})$ is convex, semi-continuous, the canonical duality relation should be replaced by the sub-differential form $\boldsymbol{\sigma} \in \partial\Phi_x(\boldsymbol{\epsilon})$, which is equivalent to

$$\boldsymbol{\sigma}^T\boldsymbol{\epsilon} = 0 \Leftrightarrow \boldsymbol{\epsilon} = \mathbf{0} \ \forall \boldsymbol{\sigma} \neq \mathbf{0}. \quad (44)$$

Thus, the integer constraint $\boldsymbol{\epsilon} = \boldsymbol{\Lambda}_x(\mathbf{x}) = \{x_i(x_i - 1)\} = \mathbf{0}$ can be relaxed by the canonical dual constraint $\boldsymbol{\sigma} \neq \mathbf{0}$ in continuous space.

The canonical duality-triality theory has been used successfully for solving mixed integer programming problems [14, 80, 83]. Particularly, for the quadratic integer programming problem (2), i.e.

$$\min \left\{ \Pi(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{x}^T\mathbf{f} \mid \mathbf{x} \in \{0, 1\}^n \right\},$$

⁶See http://en.wikipedia.org/wiki/Mathematical_optimization

the canonical dual is [26, 68]

$$\max \left\{ \Pi^d(\boldsymbol{\sigma}) = -\frac{1}{2}(\mathbf{f} + \boldsymbol{\sigma})^T \mathbf{G}^{-1}(\boldsymbol{\sigma})(\mathbf{f} + \boldsymbol{\sigma}) \mid \boldsymbol{\sigma} \in \mathcal{S}_c^+ \right\} \quad (45)$$

where $\mathbf{G}(\boldsymbol{\sigma}) = \mathbf{Q} + 2\text{Diag}(\boldsymbol{\sigma})$. This is a concave maximization problem over the convex set in continuous space

$$\mathcal{S}_c^+ = \{\boldsymbol{\sigma} \in \mathbb{R}^n \mid \boldsymbol{\sigma} \neq \mathbf{0}, \mathbf{G}(\boldsymbol{\sigma}) \succ 0\},$$

which can be solved easily if $\mathcal{S}_c^+ \neq \emptyset$. Otherwise, the integer programming problem (2) could be NP-hard, which is a conjecture proposed in [68]. In this case, a second canonical dual problem has been proposed in [69, 87]

$$\min \left\{ \Pi^g(\boldsymbol{\sigma}) = -\frac{1}{2}\boldsymbol{\sigma}^T \mathbf{Q}^{-1}\boldsymbol{\sigma} - \sum_{i=1}^n |f_i - \sigma_i| \mid \boldsymbol{\sigma} \in \mathbb{R}^n \right\}. \quad (46)$$

This is a unconstrained nonsmooth minimization problem, which can be solved by some deterministic methods, such as DIRECT method [87].

Remark 2 (Subjective Function and NP-hard Problems) The subjective function $F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle$ in the general model $\Pi(\boldsymbol{\chi}) = W(D\boldsymbol{\chi}) - F(\boldsymbol{\chi})$ plays an important role in global optimization problems. It was proved in [69] that for quadratic integer programming problem (2), if the source term \mathbf{f} is bigger enough, the solution is simply $\{x_i\} = \{0 \text{ if } f_i < 0, 1 \text{ if } f_i > 0\}$ (Theorem 8, [69]). If a system has no input, by Newton's law, it has either trivial solution or infinite number solutions. For example, the well-known max-cut problem

$$\max \left\{ \Pi(\mathbf{x}) = \frac{1}{4} \sum_{i,j=1}^{n+1} \omega_{ij}(1 - x_i x_j) \mid x_i \in \{-1, 1\} \forall i = 1, \dots, n \right\} \quad (47)$$

is a special case of quadratic integer programming problem without the linear term. The integer condition is a physical (constitutive) constraint. Since there is no geometrical constraint, the graph is not fixed and any rigid motion is possible. Due to the symmetry $\omega_{ij} = \omega_{ji} > 0$, the global solution is not unique. The canonical dual feasible space \mathcal{S}_c^+ in this example is empty and the problem is considered as NP-complete even if $\omega_{ij} = 1$ for all edges $i, j = 1, \dots, n$ [100]. However, by adding a linear perturbation term, this problem can be solved efficiently by the canonical duality theory [155].

3.2 Unified model in mathematical physics

In analysis and mathematical physics, the configuration variable $\boldsymbol{\chi}(t, \mathbf{x})$ is a continuous field function $\boldsymbol{\chi} : [0, T] \times \Omega \subset \mathbb{R} \times \mathbb{R}^d \rightarrow \omega \subset \mathbb{R}^p$ (which is a hyper-surface if $d + 1 = p$ in differential geometry). The linear operator $D = (\partial_t, \partial_x)$ is a partial differential operator and the stored energy $W(D\boldsymbol{\chi}) = T(\partial_t \boldsymbol{\chi}) - U(\partial_x \boldsymbol{\chi})$ with $T(\mathbf{v})$ as the kinetic energy and $U(\boldsymbol{\epsilon})$ as deformation energy. Since $\mathbf{v} = \partial_t \boldsymbol{\chi}$ is a vector, the objectivity for kinetic energy $T(\mathbf{v})$ is also known as isotropy. But $\boldsymbol{\epsilon} = \partial_x \boldsymbol{\chi}$ is a tensor, the deformation energy $U(\boldsymbol{\epsilon})$ should be an objective function. In this case, the Gao and Strang model (3) is

$$\min \{ \Pi(\boldsymbol{\chi}) = T(\partial_t \boldsymbol{\chi}) - U(\partial_x \boldsymbol{\chi}) - \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle \mid \boldsymbol{\chi} \in \mathcal{X}_c \}. \quad (48)$$

The stationary condition $\delta\Pi(\boldsymbol{\chi}) = 0$ leads to a general nonlinear partial differential equation

$$\partial_t^* \partial_{\mathbf{v}} T(\partial_t \boldsymbol{\chi}) - \partial_x^* \partial_{\boldsymbol{\epsilon}} U(\partial_x \boldsymbol{\chi}) = \bar{\boldsymbol{\chi}}^*. \quad (49)$$

The nonlinearity of this equation mainly depends on T and U . For Newtonian mechanics, $T(\mathbf{v})$ is quadratic. By the objectivity, the deformation energy $U(\boldsymbol{\epsilon})$ can also be split into quadratic part and a nonlinear part such that $W(D\boldsymbol{\chi}) = \frac{1}{2} \langle \boldsymbol{\chi}, \mathbf{Q}\boldsymbol{\chi} \rangle + V(\mathbf{D}\boldsymbol{\chi})$, where $\mathbf{Q} : \mathcal{X}_c \rightarrow \mathcal{X}^*$ is a self-adjoint operator, \mathbf{D} is a linear operator, and $V(\boldsymbol{\epsilon})$ is a nonlinear objective functional. The most simple example is a fourth-order polynomial

$$V(\boldsymbol{\epsilon}) = \int_{\Omega} \frac{1}{2} \left(\frac{1}{2} \|\boldsymbol{\epsilon}\|^2 - \lambda \right)^2 d\Omega, \quad (50)$$

which is nonconvex for $\lambda > 0$. This nonconvex functional appears extensively in mathematical physics. In fluid mechanics and thermodynamics, $V(\boldsymbol{\epsilon})$ is the well-known *van de Waals double-well energy*. It is also known as the *sombrero potential* in cosmic string theory [20], or the *Mexican hat* in *Higgs mechanism* [22] and quantum field theory [99]. For this most simple nonconvex potential, the general model (3) can be written as

$$\mathbf{Q}\boldsymbol{\chi} + \mathbf{D}^* \left[\left(\frac{1}{2} \|\mathbf{D}\boldsymbol{\chi}\|^2 - \lambda \right) \mathbf{D}\boldsymbol{\chi} \right] = \bar{\boldsymbol{\chi}}^*. \quad (51)$$

This model covers many well-known equations.

1) **Duffing equation** ($\mathbf{Q} = -\partial_t^2$ and $\mathbf{D} = \mathbf{I}$ is an identical operator):

$$\chi_{tt} + \left(\frac{1}{2} \chi^2 - \lambda \right) \chi = f(t) \quad (52)$$

2) **Landau-Ginzburg equation** ($\mathbf{Q} = -\Delta$, $\mathbf{D} = \mathbf{I}$):

$$-\Delta \chi + \left(\frac{1}{2} \|\chi\|^2 - \lambda \right) \chi = \mathbf{f} \quad (53)$$

3) **Cahn-Hilliar equation** ($\mathbf{Q} = -\Delta + \text{curlcurl}$, $\mathbf{D} = \mathbf{I}$):

$$-\Delta \chi + \text{curlcurl} \chi + \left(\frac{1}{2} \|\chi\|^2 - \lambda \right) \chi = \mathbf{f}. \quad (54)$$

4) **Nonlinear Gorden equation** ($\mathbf{Q} = -\partial_{tt} + \Delta$, $\mathbf{D} = \mathbf{I}$):

$$-\chi_{tt} + \Delta \chi + \left(\frac{1}{2} \|\chi\|^2 - \lambda \right) \chi = \mathbf{f}. \quad (55)$$

5) **Nonlinear Gao beam** ($\mathbf{Q} = \rho \partial_{tt} + K \partial_{xxxx}$, $\mathbf{D} = \partial_x$):

$$\rho \chi_{tt} + K \chi_{xxxx} - \left[\left(\frac{1}{2} \chi_x^2 - \lambda \right) \chi_x \right]_x = f, \quad (56)$$

where $\lambda \in \mathbb{R}$ is an axial force and $f(t, x)$ is the lateral load.

According to the nonlinear classification discussed in Section 2.2, the general equation (51) is semilinear as long as $\mathbf{D} = \mathbf{I}$. While the nonlinear Gao beam is quasi-linear.

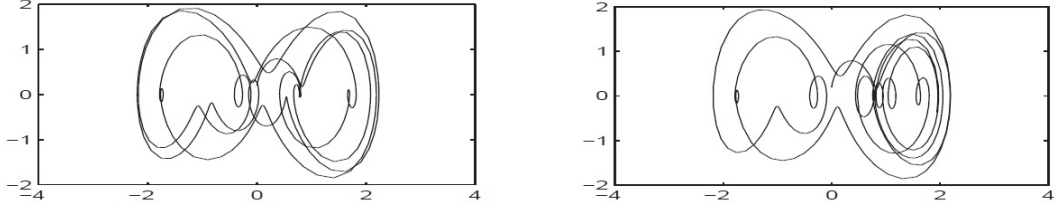


Figure 2: Chaotical trajectories of the nonlinear Gao beam computed by “ode23” (left) and “ode15s” (right) in MATLAB

However, if $\lambda > 0$, all these PDEs equations are geometrically nonlinear but physically linear since by the canonical transformation

$$\boldsymbol{\xi} = \Lambda(\boldsymbol{\epsilon}) = \frac{1}{2}\|\boldsymbol{\epsilon}\|^2 - \lambda, \quad V(\boldsymbol{\epsilon}) = \Phi(\Lambda(\boldsymbol{\epsilon})) = \int_{\Omega} \frac{1}{2}\boldsymbol{\xi}^2 d\Omega,$$

the canonical duality relation $\boldsymbol{\xi}^* = \partial\Phi(\boldsymbol{\xi}) = \boldsymbol{\xi}$ is linear.

The geometrical nonlinearity represents large deformation in continuum physics, or far from the equilibrium state in complex systems, which is necessary for nonconvexity but not sufficient. The nonconvexity of a geometrically nonlinear problem depends on external force and internal parameters. For example, the total potential of the nonlinear Gao beam is nonconvex only if the compressive load $\lambda > \lambda_c$, the Euler buckling load, i.e. the first eigenvalue of $K\chi_{xxxx}$ [43, 52]. In this case, the two minimizers represent the two buckled states, while the local maximizer represents the unbuckled (unstable) state. For dynamical loading, these two local minimizers are very sensitive to the driving force and initial conditions this nonconvex beam model could produce chaotic vibration. The so-called strange attractor is actually a local minimizer [60, 61]. Particularly, if the variable $\chi(t, x)$ can be separate variable as $\chi = q(t) \sin(\theta x)$, this nonlinear beam model is equivalent to the Duffing equation, which is well-known in chaotic dynamics. Figure 2 shows clearly that for the same given initial data, the same Runger-Kutta iteration but with different solvers in MATLAB produces very different “trajectories” in phase space $q-p$ ($p = q_t$). Therefore, this nonlinear beam model is important for understanding many challenging problems in both mathematics and engineering applications and has been subjected to extensive study recently [1, 10, 13, 60, 102, 113, 117].

The canonical duality theory has been successfully for modeling real-world problems in nonconvex/nonsmooth dynamical systems [59], differential geometry [90], contact mechanics [48], post-buckling structures [52], multi-scale phase transitions of solids [92], and general mathematical physics (see Chapter 4, [53]).

4 Applications in large deformation mechanics

For mixed boundary-value problems, the input $\bar{\boldsymbol{\chi}}^*$ is the body force \mathbf{f} in the domain $\Omega \subset \mathbb{R}^d$ and surface traction \mathbf{t} on the boundary $\Gamma_t \subset \partial\Omega$. The external energy

$$F(\boldsymbol{\chi}) = \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle = \int_{\Omega} \boldsymbol{\chi} \cdot \mathbf{f} d\Omega + \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} d\Gamma \quad (57)$$

is a linear functional defined on $\mathcal{X}_a = \{\boldsymbol{\chi} \in \mathcal{C}^1[\Omega; \mathbb{R}^p] \mid \boldsymbol{\chi} = 0 \text{ on } \Gamma_{\chi}\}$. For a hyper-elastic material deformation problem, we have $\dim \Omega = d = p = 3$. The stored energy

$W(\mathbf{F})$ is usually a nonconvex functional of the deformation gradient tensor $\mathbf{F} = \nabla \boldsymbol{\chi}$

$$W(\mathbf{F}) = \int_{\Omega} U(\mathbf{F}) \, d\Omega, \quad (58)$$

where $U(\mathbf{F})$ is the stored energy density defined on $\mathcal{W}_a = \mathbb{M}_+^3 = \{\mathbf{F} = \{F_{\alpha}^i\} \in \mathbb{R}^{3 \times 3} \mid \det \mathbf{F} > 0\}$. Thus, on the kinetically admissible space

$$\mathcal{X}_c = \{\boldsymbol{\chi} \in \mathcal{C}^1[\Omega; \mathbb{R}^d] \mid \det(\nabla \boldsymbol{\chi}) > 0, \boldsymbol{\chi} = 0 \text{ on } \Gamma_{\chi}\},$$

the general model (3) is a typical nonconvex variational problem

$$\min_{\boldsymbol{\chi} \in \mathcal{X}_c} \left\{ \Pi(\boldsymbol{\chi}) = \int_{\Omega} U(\nabla \boldsymbol{\chi}) \, d\Omega - \int_{\Omega} \boldsymbol{\chi} \cdot \mathbf{f} \, d\Omega - \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} \, d\Gamma \right\}. \quad (59)$$

The linear operator $D = \text{grad} : \mathcal{X}_a \rightarrow \mathbb{M}_+^3$ in this problem is a gradient. The stationary condition $\delta \Pi(\boldsymbol{\chi}) = 0$ leads to a mixed boundary-value problem (BVP)

$$(BVP) : \quad A(\boldsymbol{\chi}) = \nabla^* \partial_{\mathbf{F}} W(\nabla \boldsymbol{\chi}) = \begin{cases} -\nabla \cdot \nabla_{\mathbf{F}} U(\nabla \boldsymbol{\chi}) = \mathbf{f} & \text{in } \Omega, \\ \mathbf{n} \cdot \nabla_{\mathbf{F}} U(\nabla \boldsymbol{\chi}) = \mathbf{t} & \text{on } \Gamma_t. \end{cases} \quad (60)$$

According to the definition of nonlinear PDEs, the first equilibrium equation (60) is fully nonlinear as long as $\partial U(\mathbf{F})$ is nonlinear. However, it is geometrically linear if $U(\mathbf{F})$ is convex. It is completely nonlinear only if $U(\mathbf{F})$ is nonconvex. Therefore, the definition of fully nonlinearity in PDEs can't be used to identify difficulty of the nonlinear problems.

It is well-known in finite deformation theory that the convexity of the stored energy density $U(\mathbf{F})$ contradicts the most immediate physical experience (see Theorem 4.8-1, [18]). Indeed, even its domain \mathbb{M}_+^3 is not a convex subset of $\mathbb{R}^{3 \times 3}$ (Theorem 4.7-4, [18]). Therefore, the solution to the (BVP) is only a stationary point of the total potential $\Pi(\boldsymbol{\chi})$. In order to identify minimizer of the problem, many generalized convexities have been suggested and the following results are well-known (see [53]):

$$U(\mathbf{F}) \text{ is convex} \Rightarrow \text{poly-convex} \Rightarrow \text{quasi-convex}^7 \Rightarrow \text{rank-one convex}. \quad (61)$$

If $U \in \mathcal{C}^2(\mathbb{M}_+^3)$, then the rank-one convexity is equivalent to the Legendre-Hadamard (L.H.) condition:

$$\sum_{i,j=1}^3 \sum_{\alpha,\beta=1}^3 \frac{\partial^2 U(\mathbf{F})}{\partial F_{\alpha}^i \partial F_{\beta}^j} a_i a_j b^{\alpha} b^{\beta} \geq 0 \quad \forall \mathbf{a} = \{a_i\} \in \mathbb{R}^3, \forall \mathbf{b} = \{b^{\alpha}\} \in \mathbb{R}^3. \quad (62)$$

The Legendre-Hadamard condition in finite elasticity is also referred to as the *ellipticity condition*, i.e., if the L.H. condition holds, the partial differential operator $A(\boldsymbol{\chi})$ in (60) is considered to be elliptic. For one-dimensional problems $\Omega \subset \mathbb{R}$, all these convexities are equivalent and the rank-one convexity is the well-known convexity in vector space. We should emphasize that these generalized convexities and L.H. condition are local criteria not global. As long as the total potential $\Pi(\boldsymbol{\chi})$ is locally nonconvex in certain domain of Ω , the boundary-value problem (60) could have multiple solutions $\boldsymbol{\chi}(\mathbf{x})$ at

⁷The quasiconvexity used in variational calculus and continuum physics has an entirely different meaning from that used in optimization, where a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called quasiconvex if its level set $\mathcal{L}_{\alpha}[f] = \{x \in \mathbb{R}^n \mid f(x) \leq \alpha\}$ is convex. For example, the nonconvex function $f(x) = \sqrt{|x|}$ is quasiconvex.

each material point $\mathbf{x} \in \Omega$ and the total potential $\Pi(\boldsymbol{\chi})$ could have infinitely number of local minimizers (see [75]). This is the main difference between nonconvex analysis and nonlinear PDEs, which is a key point to understand NP-hard problems in computer science and global optimization. Unfortunately, this difference is not fully understood in both fields. It turns out that extensive efforts have been devoted for solving nonconvex variational problems directly. It was discovered by Gao and Ogden in 2008 that even for one-dimensional problems, the L.H. condition can only identify local local minimizers, and a geometrically nonlinear ODE could have infinite number solutions, both local and global minimal solutions could be nonsmooth and can't be determined by any Newton type of numerical methods [75].

By the objectivity of the stored energy density $U(\mathbf{F})$, it is reasonable to assume a canonical function $V(\mathbf{C})$ such that the following canonical transformation holds:

$$W(\mathbf{F}) = \Phi(\Lambda(\mathbf{F})) = \int_{\Omega} V(\mathbf{F}^T \mathbf{F}) d\Omega. \quad (63)$$

In this transformation, the geometrical nonlinear operator $\Lambda(\mathbf{F}) = \mathbf{F}^T \mathbf{F}$ is quadratic (objective) and $\mathbf{C} = \mathbf{F}^T \mathbf{F} \in \mathbb{S}^+ = \{\mathbf{C} = \{C_{\alpha\beta}\} \in \mathbb{R}^{3 \times 3} \mid \mathbf{C} = \mathbf{C}^T, \mathbf{C} \succ 0\}$ is the well-known right Cauchy-Green strain tensor. Its canonical dual $\mathbf{S} = \partial\Phi(\mathbf{C}) = \nabla V(\mathbf{C}) \in \mathbb{S}$ is a *second Piola-Kirchhoff type stress tensor*⁸. In terms of the canonical strain measure $\mathbf{C}(\mathbf{F})$, the kinetically admissible space $\mathcal{X}_c = \{\boldsymbol{\chi} \in \mathcal{C}^1[\Omega, \mathbb{R}^3] \mid \mathbf{C}(\nabla \boldsymbol{\chi}) \in \mathbb{S}^+, \boldsymbol{\chi} = 0 \text{ on } \Gamma_{\chi}\}$ is convex and the nonconvex variational problem (59) can be written in the canonical form

$$\min \{\Pi(\boldsymbol{\chi}) = \Phi(\mathbf{C}(\nabla \boldsymbol{\chi})) - \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle \mid \boldsymbol{\chi} \in \mathcal{X}_c\}. \quad (64)$$

By the Legendre transformation $V^*(\mathbf{S}) = \{\mathbf{C} : \mathbf{S} = \nabla V(\mathbf{C})\}$, the total complementary functional $\Xi(\boldsymbol{\chi}, \mathbf{S})$ has the following form:

$$\Xi(\boldsymbol{\chi}, \mathbf{S}) = \int_{\Omega} [\mathbf{C}(\nabla \boldsymbol{\chi}) : \mathbf{S} - V^*(\mathbf{S}) - \boldsymbol{\chi} \cdot \mathbf{f}] d\Omega - \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} d\Gamma. \quad (65)$$

By the fact that the linear operator $D = \text{grad}$ is a differential operator, it is difficult to find its inverse operator. In order to obtain the canonical dual $\Pi^d(\mathbf{T})$, we need to introduce the following *statically admissible space*

$$\mathcal{T}_c = \{\boldsymbol{\tau} \in \mathcal{C}^1[\Omega; \mathbb{R}^{3 \times 3}] \mid -\nabla \cdot \boldsymbol{\tau} = \mathbf{f} \text{ in } \Omega, \mathbf{n} \cdot \boldsymbol{\tau} = \mathbf{t} \text{ on } \Gamma_t\}.$$

Clearly, for any given $\boldsymbol{\chi} \in \mathcal{X}_a = \{\boldsymbol{\chi} \in \mathcal{C}^1[\Omega; \mathbb{R}^3] \mid \det(\nabla \boldsymbol{\chi}) > 0, \boldsymbol{\chi} = 0 \text{ on } \Gamma_{\chi}\}$, the external energy $F(\boldsymbol{\chi})$ can be written equivalently as

$$F_{\boldsymbol{\tau}}(\boldsymbol{\chi}) = \int_{\Omega} \boldsymbol{\chi} \cdot (-\nabla \cdot \boldsymbol{\tau}) d\Omega + \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} d\Gamma = \int_{\Omega} (\nabla \boldsymbol{\chi}) : \boldsymbol{\tau} d\Omega \quad \forall \boldsymbol{\tau} \in \mathcal{T}_c \quad (66)$$

Thus, for any given $\boldsymbol{\tau} \in \mathcal{T}_c$, the Λ -conjugate of $F(\boldsymbol{\chi})$ can be obtained

$$F_{\boldsymbol{\tau}}^{\Lambda}(\mathbf{S}) = \text{sta}\{\langle \mathbf{C}(\nabla \boldsymbol{\chi}); \mathbf{S} \rangle - F_{\boldsymbol{\tau}}(\boldsymbol{\chi}) \mid \boldsymbol{\chi} \in \mathcal{X}_a\} = - \int_{\Omega} \frac{1}{4} \text{tr}(\boldsymbol{\tau} \cdot \mathbf{S}^{-1} \cdot \boldsymbol{\tau}^T) d\Omega. \quad (67)$$

Its domain should be

$$\mathcal{S}_c = \{\mathbf{S} \in \mathcal{E}_a^* \mid \det(\boldsymbol{\tau} \cdot \mathbf{S}^{-1}) > 0\}. \quad (68)$$

⁸The second Piola-Kirchhoff stress tensor is defined by $\mathbf{T} = \partial\Phi(\mathbf{E})$, where $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I})$ is the Green-St. Venant strain tensor. Therefore, we have $\mathbf{S} = 2\mathbf{T}$.

Therefore, the pure complementary energy can be obtained as

$$\Pi^d(\mathbf{S}; \boldsymbol{\tau}) = - \int_{\Omega} \left[\frac{1}{4} \text{tr}(\boldsymbol{\tau} \cdot \mathbf{S}^{-1} \cdot \boldsymbol{\tau}^T) + V^*(\mathbf{S}) \right] d\Omega, \quad (69)$$

which depends on not only the canonical stress $\mathbf{S} \in \mathcal{S}_c$, but also the statically admissible field $\boldsymbol{\tau} \in \mathcal{T}_c$. Let

$$\mathcal{S}_c^+ = \{\mathbf{S} \in \mathcal{S}_c \mid \mathbf{S} \succ 0\}, \quad \mathcal{S}_c^- = \{\mathbf{S} \in \mathcal{S}_c \mid \mathbf{S} \prec 0\}. \quad (70)$$

Theorem 4 (Pure Complementary Energy Principle, Gao [47, 49, 53])

If $(\bar{\mathbf{S}}, \bar{\boldsymbol{\tau}}) \in \mathcal{S}_c \times \mathcal{T}_c$ is a stationary points of $\Pi^d(\mathbf{S}; \boldsymbol{\tau})$, then the deformation defined by

$$\bar{\boldsymbol{\chi}}(\mathbf{x}) = \frac{1}{2} \int_{\mathbf{x}_0}^{\mathbf{x}} \bar{\boldsymbol{\tau}} \cdot \bar{\mathbf{S}}^{-1} d\mathbf{x} \quad (71)$$

along any path from $\mathbf{x}_0 \in \Gamma_{\chi}$ to $\mathbf{x} \in \Omega$ is a critical point of $\Pi(\boldsymbol{\chi})$ and $\Pi(\bar{\boldsymbol{\chi}}) = \Pi^d(\bar{\mathbf{S}}; \bar{\boldsymbol{\tau}})$. Moreover, $\bar{\boldsymbol{\chi}}(\mathbf{x})$ is a global minimizer of $\Pi(\boldsymbol{\chi})$ if $\bar{\mathbf{S}}(\mathbf{x}) \in \mathcal{S}_c^+ \quad \forall \mathbf{x} \in \Omega$.

The vector-valued function $\bar{\boldsymbol{\chi}}(\mathbf{x})$ is a solution to the boundary-value problem of the second equilibrium equation in (60) if the compatibility condition $\nabla \times (\bar{\boldsymbol{\tau}} \cdot \bar{\mathbf{S}}^{-1}) = 0$ holds.

Proof. Using Lagrange multiplier $\boldsymbol{\chi} \in \mathcal{X}_a$ to relax the equilibrium conditions in \mathcal{T}_c , we have

$$\Theta(\mathbf{S}; \boldsymbol{\tau}, \boldsymbol{\chi}) = - \int_{\Omega} \left[\frac{1}{4} \text{tr}(\boldsymbol{\tau} \cdot \mathbf{S}^{-1} \cdot \boldsymbol{\tau}^T) + V^*(\mathbf{S}) \right] d\Omega - \int_{\Omega} \boldsymbol{\chi} \cdot (\nabla \cdot \boldsymbol{\tau} + \mathbf{f}) d\Omega + \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} d\Gamma. \quad (72)$$

Its stationary condition leads to

$$2\nabla \boldsymbol{\chi} = \boldsymbol{\tau} \cdot \mathbf{S}^{-1} \quad (73)$$

$$4\mathbf{S} \cdot (\nabla V^*(\mathbf{S})) \cdot \mathbf{S} = \boldsymbol{\tau}^T \cdot \boldsymbol{\tau} \quad (74)$$

and the equilibrium equations in \mathcal{T}_c . From (73) we have $\boldsymbol{\tau} = 2(\nabla \boldsymbol{\chi}) \cdot \mathbf{S}$. Substituting this into (74) we have $(\nabla \boldsymbol{\chi})^T (\nabla \boldsymbol{\chi}) = \nabla V^*(\mathbf{S})$, which is equivalent to $\mathbf{S} = \nabla V(\mathbf{C}(\nabla \boldsymbol{\chi}))$ due to the canonical duality. Thus, from the canonical transformation, we have

$$\boldsymbol{\tau} = 2(\nabla \boldsymbol{\chi}) \cdot (\nabla_{\mathbf{C}} V(\mathbf{C}(\nabla \boldsymbol{\chi}))) = \nabla_{\mathbf{F}} U(\nabla \boldsymbol{\chi}) \quad (75)$$

due to the chain rule. This shows that the integral (71) is indeed a stationary point of $\Pi(\boldsymbol{\chi})$ since $\boldsymbol{\tau} \in \mathcal{T}_c$.

By the fact that $\mathbf{C} = \Lambda(\mathbf{F})$ is a quadratic operator, the Gao-Strang gap function is

$$G_{ap}(\boldsymbol{\chi}, \mathbf{S}) = \int_{\Omega} \text{tr}[(\nabla \boldsymbol{\chi}) \cdot \mathbf{S} \cdot (\nabla \boldsymbol{\chi})] d\Omega.$$

Clearly, $G_{ap}(\boldsymbol{\chi}, \mathbf{S})$ is non negative for any given $\boldsymbol{\chi} \in \mathcal{X}_a$ if and only if $\mathbf{S}(\mathbf{x}) \in \mathcal{S}_c^+ \quad \forall \mathbf{x} \in \Omega$. Replace $\nabla \boldsymbol{\chi} = \frac{1}{2} \boldsymbol{\tau} \cdot \mathbf{S}^{-1}$, this gap function reads

$$G_{ap}(\boldsymbol{\chi}(\mathbf{S}, \boldsymbol{\tau}), \mathbf{S}) = \int_{\Omega} \frac{1}{4} \text{tr}[\boldsymbol{\tau} \cdot \mathbf{S}^{-1} \cdot \boldsymbol{\tau}^T] d\Omega,$$

which is convex for any $\boldsymbol{\tau} \in \mathcal{T}_c$ if and only if $\mathbf{S}(\mathbf{x}) \in \mathcal{S}_c^+ \quad \forall \mathbf{x} \in \Omega$. Therefore, the canonical dual $\Pi^d(\mathbf{S}; \boldsymbol{\tau})$ is concave on $\mathcal{S}_c^+ \times \mathcal{T}_c$. By the canonical min-max duality, $\bar{\boldsymbol{\chi}}$ is a unique global minimizer if $\bar{\mathbf{S}}(\mathbf{x}) \succ 0 \quad \forall \mathbf{x} \in \Omega$.

The compatibility condition $\nabla \times (\boldsymbol{\tau} \cdot \mathbf{S}) = 0$ is necessary for an analytical solution to the mixed boundary-value problem (60) due to the fact that $\text{curl grad } \boldsymbol{\chi} = 0$. \square

The pure complementary energy principle was first proposed by Gao (1997) in post-buckling problems of a large deformed beam [46]. Generalization to 3-D finite deformation theory and nonconvex analysis were given during 1998-2000 [47, 49, 51, 53, 54]. The equation (74) is called the *canonical dual algebraic equation* first obtained in 1998 [47]. This equation shows that by the canonical dual transformation, the nonlinear partial differential equation can be equivalently reformed as an algebraic equation. The equation (75) show that the statically admissible field $\boldsymbol{\tau} = \nabla U(\mathbf{F})$ is actually the *first Piola-Kirchhoff stress*. For one-dimensional problems, $\boldsymbol{\tau} \in \mathcal{T}_c$ can be easily obtained by the given input. For geometrically nonlinear problems, $\nabla V^*(\mathbf{S})$ is linear and (74) can be solved analytically to obtain a complete set of analytical solutions [53, 54, 70, 75, 76]. By the triality theory, the positive solution $\mathbf{S} \in \mathcal{S}_c^+$ produces a global minimal solution $\bar{\boldsymbol{\chi}}$, while the negative $\mathbf{S} \in \mathcal{S}_c^-$ can be used to identify local extremal solutions. To see this, let us consider the Hessian of the stored energy $U(\mathbf{F}) = V(\mathbf{C}(\mathbf{F}))$. By chain rule, we have

$$\frac{\partial^2 U(\mathbf{F})}{\partial F_\alpha^i \partial F_\beta^j} = 2\delta^{ij} S_{\alpha\beta} + 4 \sum_{\theta, \nu=1}^3 F_\theta^i H_{\theta\alpha\beta\nu} F_\nu^j, \quad (76)$$

where $\mathbf{H} = \{H_{\theta\alpha\beta\nu}\} = \nabla^2 V(\mathbf{C}) \succ 0$ due to the convexity of the canonical function $V(\mathbf{C})$. Clearly, if $\mathbf{S} \succeq 0$, the L.H. condition holds and the associated $\bar{\boldsymbol{\chi}}$ is a global minimal solution. By the fact that $2\mathbf{F} = \boldsymbol{\tau} \mathbf{S}^{-1}$, we know that $\nabla^2 U(\mathbf{F})$ could be either positive or negative definite even if $\mathbf{S} \prec 0$. Therefore, depending the eigenvalues of $\mathbf{S} \prec 0$, the L.H. condition could also hold at a local minimizer of $\Pi(\boldsymbol{\chi})$ [70]. This shows that the triality theory can be used to identify both global and local extremal solutions, while the L.H. condition is only a necessary condition for a local minimal solution. It is known that an elliptic equation is corresponding to a convex variational problem. Therefore, it is a question if the Legendre-Hadamard condition can still be called as the ellipticity condition in finite elasticity and nonconvex analysis. By the fact that the well-known open problem left by Reissner *et al* [137] has been solved by Theorem 4, the pure complementary energy principle is known as the Gao principle in literature (see [114]).

The canonical transformation $W(\mathbf{F}) = \Phi(\Lambda(\mathbf{F}))$ is not unique since the geometrical operator $\Lambda(\mathbf{F})$ can be chosen differently to have different canonical strain measures. For example, the well-known *Hill-Seth strain family*

$$\mathbf{E}^{(\eta)} = \Lambda(\mathbf{F}) = \frac{1}{2\eta} [(\mathbf{F}^T \cdot \mathbf{F})^\eta - \mathbf{I}] \quad (77)$$

is a geometrically admissible objective strain measure for any given $\eta \in \mathbb{R}$ (see Definition 6.3.1, [53]). Particularly, $\mathbf{E}^{(1)}$ is the well-known *Green-St. Venant strain tensor* \mathbf{E} . For *St. Venant-Kirchhoff materials*, the stored strain density is quadratic: $V(\mathbf{E}) = \frac{1}{2} \mathbf{E} : \mathbf{H} : \mathbf{E}$, where \mathbf{H} is the Hooke tensor. Clearly, $V(\mathbf{E})$ is convex but

$$U(\mathbf{F}) = V(\mathbf{E}(\mathbf{F})) = \frac{1}{8} (\mathbf{F}^T \cdot \mathbf{F} - \mathbf{I}) : \mathbf{H} : (\mathbf{F}^T \cdot \mathbf{F} - \mathbf{I})$$

is a (nonconvex) double-well type function of \mathbf{F} , which is not even rank-one convex [136]. The canonical duality is linear $\mathbf{T} = \nabla V(\mathbf{E}) = \mathbf{H} : \mathbf{E}$ and the generalized total complementary energy $\Xi(\boldsymbol{\chi}, \mathbf{T})$ is the well-known Hellinger-Reissner complementary energy

$$\Xi(\boldsymbol{\chi}, \mathbf{T}) = \int_{\Omega} \left[\mathbf{E}(\nabla \boldsymbol{\chi}) : \mathbf{T} - \frac{1}{2} \mathbf{T} : \mathbf{H}^{-1} : \mathbf{T} - \boldsymbol{\chi} \cdot \mathbf{f} \right] d\Omega - \int_{\Gamma_t} \boldsymbol{\chi} \cdot \mathbf{t} d\Gamma. \quad (78)$$

In this case, the primal problem (59) is a geometrically nonlinear variational problem, and its canonical dual functional is

$$\Pi^d(\mathbf{T}; \boldsymbol{\tau}) = - \int_{\Omega} \frac{1}{2} [\text{tr}(\boldsymbol{\tau} \cdot \mathbf{T}^{-1} \cdot \boldsymbol{\tau}^T + \mathbf{T}) + \mathbf{T} : \mathbf{H}^{-1} : \mathbf{T}] d\Omega. \quad (79)$$

The canonical dual algebraic equation (74) is then a cubic tensor equation

$$2 \mathbf{T} \cdot (\mathbf{H}^{-1} : \mathbf{T} + \mathbf{I}) \cdot \mathbf{T} = \boldsymbol{\tau}^T \cdot \boldsymbol{\tau} \quad (80)$$

For a given statically admissible stress field $\boldsymbol{\tau} \in \mathcal{T}_c$, this tensor equation could have at most 27 solutions $\mathbf{T}(\mathbf{x})$ at each material point $\mathbf{x} \in \Omega$, but only one $\mathbf{T}(\mathbf{x}) \succ 0$, which leads to a global minimal solution [72].

For many real-world problems, the statically admissible stress $\boldsymbol{\tau} \in \mathcal{T}_c$ can be uniquely obtained and the canonical dual algebraic equation (80) can be solved to obtain all possible stress solutions. The canonical duality-triality theory has been used successfully for solving a class of nonconvex variational/boundary value problems [54, 73, 75], pure azimuthal shear [76] and anti-plane shear problems [70].

5 Applications to computational mechanics and global optimization

Numericalization for solving the nonconvex variational problem (3) leads to a global optimization problem in a finite dimensional space $\mathcal{X} = \mathcal{X}^*$. In complex systems, the decision variable $\boldsymbol{\chi}$ could be either vector or matrix. In operations research, such as logistic and supply chain management sciences, $\boldsymbol{\chi}$ can be even a high-order matrix $\boldsymbol{\chi} = \{\chi_{ij\dots k}\}$. Correspondingly, the linear operator $D : \mathcal{X}_a \rightarrow \mathcal{W}_a$ is a matrix or high-order tensor. In general global optimization problems, the internal energy $W(D\boldsymbol{\chi})$ is not necessary to be an objective function. As long as the canonical transformation $W(D\boldsymbol{\chi}) = \Phi(\Lambda(D\boldsymbol{\chi}))$ holds, the canonical duality-triality theory can be used for solving a large class of nonconvex/discrete optimization problems.

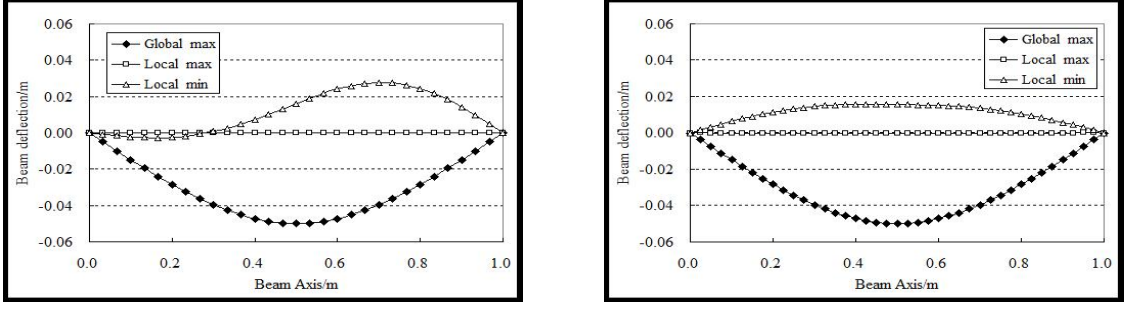
5.1 Canonical dual finite element method

It was shown in [44] that by using independent finite element interpolations for displacement and generalized stress:

$$\boldsymbol{\chi}(\mathbf{x}) = \mathbf{N}_u(\mathbf{x})\mathbf{q}^e, \quad \mathbf{S}(\mathbf{x}) = \mathbf{N}_\varsigma(\mathbf{x})\mathbf{p}^e \quad \forall \mathbf{x} \in \Omega^e \subset \Omega, \quad (81)$$

the total complementary functional $\Xi(\boldsymbol{\chi}, \mathbf{S})$ defined by (65) can be discretized as a function in finite dimensional space

$$\Xi(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \mathbf{q}^T \mathbf{G}(\mathbf{p}) \mathbf{q} - \Phi^*(\mathbf{p}) - \mathbf{q}^T \mathbf{f}, \quad (82)$$



(a) 30 elements.

(b) 40 elements.

Figure 3: Canonical dual FEM solutions for post-buckled nonlinear beam: Global minimal solution, i.e. stable buckled state (dotted); local min, i.e. unstable buckled state (triangle); and local max, i.e. unbuckled state (squared).

where \mathbf{f} is the generalized force and $\mathbf{G}(\mathbf{p})$ is the Hessian matrix of the discretized Gao-Strang complementary gap function. In this case, the pure complementary energy can be formulated explicitly as [44]

$$\Pi^d(\mathbf{p}) = -\frac{1}{2}\mathbf{f}^T \mathbf{G}^+(\mathbf{p})\mathbf{f} - \Phi^*(\mathbf{p}), \quad (83)$$

where \mathbf{G}^+ represents a generalized inverse of \mathbf{G} . Let

$$\mathcal{S}_c^+ = \{\mathbf{p} \in \mathbb{R}^m \mid \mathbf{G}(\mathbf{p}) \succeq 0\}, \quad \mathcal{S}_c^- = \{\mathbf{p} \in \mathbb{R}^m \mid \mathbf{G}(\mathbf{p}) \prec 0\}.$$

By the fact that $\Pi^d(\mathbf{p})$ is concave on the convex set \mathcal{S}^+ , the canonical dual FE programming problem

$$\max\{\Pi^d(\mathbf{p}) \mid \mathbf{p} \in \mathcal{S}_c^+\} \quad (84)$$

can be solved easily (if $\mathcal{S}_c^+ \neq \emptyset$) to obtain the global maximizer $\bar{\mathbf{p}}$. By the triality theory, we know that $\bar{\mathbf{q}} = \mathbf{G}^+(\bar{\mathbf{p}})\mathbf{f}$ is a global minimizer of the nonconvex potential $\Pi(\mathbf{q})$. On the other hand, if $\dim \mathbf{q} = \dim \mathbf{p}$, the biggest local min and local max of $\Pi(\mathbf{q})$ can be obtained respectively by [88]

$$\min\{\Pi^d(\mathbf{p}) \mid \mathbf{p} \in \mathcal{S}_c^-\}, \quad \max\{\Pi^d(\mathbf{p}) \mid \mathbf{p} \in \mathcal{S}_c^-\}.$$

The canonical dual FEM has been used successfully in phase transitions of solids [92] and in post-buckling analysis for the large deformation beam model (2) to obtain all three possible solutions [143] (see Fig. 3). It was discovered that the local minimum is very sensitive to the lateral load and the size of the finite element meshes (see Fig. 3). This method can be used for solving general nonconvex mechanics problems.

5.2 Global optimal solutions for discrete nonlinear dynamical systems

General nonlinear dynamical systems can be modeled as a nonlinear initial-value problem

$$\chi'(t) = \mathbf{F}(t, \chi(t)) \quad t \in [0, T], \quad \chi(0) = \chi_0, \quad (85)$$

where $T > 0$, $\mathbf{F} : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a given vector-valued function. Generally speaking, if a nonlinear equation has multiple solutions at each time t in a subset of its domain

$[0, T]$, then the associated initial-valued problem should have infinite number of solutions since the unknown $\chi(t)$ is a continuous function. With time step size $h = T/n$, a discretization of the configuration $\chi(t)$ is $\mathbf{X} = (\chi_0, \chi_1, \dots, \chi_n) \in \mathcal{X}_a \subset \mathbb{R}^{d \times (n+1)}$. By the finite difference method, the initial value problem (85) can be written approximately as

$$\chi_k = \chi_{k-1} + \frac{1}{2}h\mathbf{F}(t_{k-1}, \chi_{k-1}), \quad k = 1, \dots, n. \quad (86)$$

This is still a nonlinear algebraic system. Clearly, any linear iteration can only produce one of the infinite number solutions, and such a numerical “solution” is very sensitive to the step-size and numerical errors. This is the reason why different numerical solvers produce totally different results, i.e. the so-called chaotic solutions. Rather than the traditional linear iteration from an initial value, we use the least squares method such that the nonlinear algebraic system (86) can be equivalently written as

$$\min_{\mathbf{X} \in \mathcal{X}_a} \left\{ \Pi(\mathbf{X}) = \frac{1}{2} \sum_{k=1}^n \|\chi_k - \chi_{k-1} - h\mathbf{F}(t_{k-1}, \chi_{k-1})\|^2 \right\}. \quad (87)$$

Clearly, for any given nonlinear function $\mathbf{F}(t, \chi(t))$, this is a global optimization problem, which could have multiple minimizers at each χ_k . Particularly, if $\mathbf{F}(t, \chi)$ is quadratic, then $\Pi(\mathbf{X})$ is a double-well typed fourth order polynomial function, and is considered to be NP-hard in global optimization even for $d = 1$ (one-dimensional systems) [4, 144]. However, by simply using the quadratic geometrical operator $\xi_k = \Lambda(\chi_k) = \mathbf{F}(t_k, \chi_k)$, the nonconvex least squares problem (87) can be solved by the canonical duality-triality theory to obtain global optimal solution. Applications have been given to the logistic map [115] and population growth problems [139].

5.3 Unconstrained nonconvex minimization

The general model (3) for unconstrained global optimization can be written in the following form

$$\min \left\{ \Pi(\chi) = W(\mathbf{D}\chi) + \frac{1}{2}\langle \chi, \mathbf{A}\chi \rangle - \langle \chi, \mathbf{f} \rangle \mid \chi \in \mathcal{X}_a \right\}, \quad (88)$$

where $\mathbf{D} : \mathcal{X}_a \rightarrow \mathcal{W}_a$ and $\mathbf{A} = \mathbf{A}^T$ are two given operators and $\mathbf{f} \in \mathcal{X}_a$ is a given input. For the nonconvex function $W(\epsilon)$, we assume that the canonical transformation $W(\mathbf{D}\chi) = \Phi(\Lambda(\chi))$ holds for a quadratic operator

$$\Lambda(\chi) = \left\{ \frac{1}{2}\chi^T \mathbf{H}_{\alpha\beta} \chi \right\} : \mathcal{X}_a \rightarrow \mathcal{E}_a \subset \mathbb{R}^{m \times m}, \quad (89)$$

where $\mathbf{H}_{\alpha\beta} = \mathbf{H}_{\alpha\beta}^T \quad \forall \alpha, \beta \in I_m = \{1, \dots, m\}$ is a linear operator such that \mathcal{E}_a is either a vector ($\beta = 1$) or tensor ($\alpha, \beta > 1$) space. By the convexity of the canonical function $\Phi : \mathcal{E}_a \rightarrow \mathbb{R}$, the canonical duality $\mathbf{S} = \partial\Phi(\xi) \in \mathcal{E}_a^* \subset \mathbb{R}^{m \times m}$ is invertible and the total complementary function $\Xi : \mathcal{X}_a \times \mathcal{E}_a^* \rightarrow \mathbb{R}$ reads

$$\Xi(\chi, \mathbf{S}) = \frac{1}{2}\langle \chi, \mathbf{G}(\mathbf{S})\chi \rangle - \Phi^*(\mathbf{S}) - \langle \chi, \mathbf{f} \rangle \quad (90)$$

where $\mathbf{G}(\mathbf{S}) = \mathbf{A} + \sum_{\alpha, \beta \in I_m} \mathbf{H}_{\alpha\beta} S_{\alpha\beta}$. Thus, on $\mathcal{S}_c^+ = \{\mathbf{S} \in \mathcal{E}_a^* \mid \mathbf{G}(\mathbf{S}) \succ 0\}$, the canonical dual problem (33) for the unconstrained global optimization reads

$$\max \left\{ \Pi^d(\mathbf{S}) = -\frac{1}{2}\langle \mathbf{G}^{-1}(\mathbf{S})\mathbf{f}, \mathbf{f} \rangle - \Phi^*(\mathbf{S}) \mid \mathbf{S} \in \mathcal{S}_c^+ \right\}. \quad (91)$$

The canonical duality-triality theory has been used successfully for solving the following nonconvex problems.

1) **Euclidian Distance Geometry Problem**

$$W(\mathbf{D}\boldsymbol{\chi}) = \sum_{i,j=1}^n \omega_{ij} [\|\boldsymbol{\chi}_i - \boldsymbol{\chi}_j\|^2 - d_{ij}]^2, \quad (92)$$

where the decision variable $\boldsymbol{\chi}_i \in \mathbb{R}^d$ is a position (location) vector, ω_{ij} , $d_{ij} > 0 \forall i, j = 1, \dots, n$, $i \neq j$ are given weight and distance parameters, respectively. The linear operator $\mathbf{D}\boldsymbol{\chi} = \{\boldsymbol{\chi}_i - \boldsymbol{\chi}_j\}$ in this problem is similar to the finite difference in numerical analysis. Such a problem appears frequently in computational biology [161], chaotic dynamics [115, 139], numerical algebra [142], sensor localization [110, 140], network communication [81], transportation optimization, as well as finite element analysis of structural mechanics [13, 92], etc. These problems are considered to be NP-hard even the Euclidian dimension $d = 1$ [4]. However, by the combination of the canonical duality-triality theory and perturbation methods, these problems can be solved efficiently (see [140]).

2) **Sum of Fractional Functions**

$$W(\mathbf{D}\boldsymbol{\chi}) = \sum_{i \in I_m} \frac{G_i(\mathbf{D}_g \boldsymbol{\chi})}{H_i(\mathbf{D}_h \boldsymbol{\chi})} \quad (93)$$

where G_i and $H_i > 0 \forall i \in I_m$ are given functions, \mathbf{D}_g and \mathbf{D}_h are linear operators.

3) **Exponential-Sum-Polynomials**

$$W(\mathbf{D}\boldsymbol{\chi}) = \sum_{i \in I_m} \exp\left(\frac{1}{2} \boldsymbol{\chi}^T \mathbf{B}_i \boldsymbol{\chi} - \alpha_i\right) + \sum_{j \in I_p} \frac{1}{2} \left(\frac{1}{2} \boldsymbol{\chi}^T \mathbf{C}_j \boldsymbol{\chi} - \beta_j\right)^2, \quad (94)$$

where \mathbf{B}_i and \mathbf{C}_j are given symmetrical matrices in $\mathbb{R}^{n \times n}$, α_i, β_j are given parameters.

4) **Log-Sum-Exp Functions**

$$W(\mathbf{D}\boldsymbol{\chi}) = \frac{1}{\beta} \log \left[1 + \sum_{i \in I_p} \exp \left(\beta \left(\frac{1}{2} \boldsymbol{\chi}^T \mathbf{B}_i \boldsymbol{\chi} + d \right) \right) \right], \quad (95)$$

where $\beta > 0$, $\mathbf{B}_i = \mathbf{B}_i^T$, and $d \in \mathbb{R}$ are given.

All these functions appear extensively in modeling real-world problems, such as computational biology [161], bio-mechanics, phase transitions [75], filter design [157], location/transportation and networks optimization [81, 140], communication and information theory (see [106]) etc. By using the canonical duality-triality theory, these problems can be solved nicely (see [17, 79, 108, 119, 163]).

5.4 Constrained global optimization

Recall the standard mathematical model in global optimization (1)

$$\min f(\mathbf{x}), \quad \text{s.t.} \quad h_i(\mathbf{x}) = 0, \quad g_j(\mathbf{x}) \leq 0 \quad \forall i \in I_m, \quad j \in I_p, \quad (96)$$

where f , g_i and h_j are differentiable, real-valued functions on a subset of \mathbb{R}^n for all $i \in I_m$ and $j \in I_p$. For notational convenience, we use vector forms for constraints

$$\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_m(\mathbf{x})), \quad \mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_p(\mathbf{x})).$$

Therefore, the feasible space can be defined as

$$\mathcal{X}_c := \{\mathbf{x} \in \mathbb{R}^n | \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{h}(\mathbf{x}) = 0\}.$$

Lagrange multiplier method was originally proposed by J-L Lagrange from analytical mechanics in 1811 [103]. During the past two hundred years, this method and the associated Lagrangian duality theory have been well-developed with extensively applications to many fields of physics, mathematics and engineering sciences. Strictly speaking, the Lagrange multiplier method can be used only for equilibrium constraints. For inequality constraints, the well-known KKT conditions are involved. Here we show that both the classical Lagrange multiplier method and the KKT theory can be unified by the canonical duality theory.

For convex constrained problem, i.e. $f(\mathbf{x})$, $\mathbf{g}(\mathbf{x})$ and $\mathbf{h}(\mathbf{x})$ are convex, the standard canonical dual transformation can be used. We can choose the geometrical operator $\boldsymbol{\xi}_0 = \boldsymbol{\Lambda}_0(\mathbf{x}) = \{\mathbf{g}(\mathbf{x}), \mathbf{h}(\mathbf{x})\} : \mathbb{R}^n \rightarrow \mathbb{R}^{m+p}$ and let

$$\Phi_0(\boldsymbol{\xi}_0) = \Psi_g(\mathbf{g}) + \Psi_h(\mathbf{h}),$$

where

$$\Psi_g(\mathbf{g}) = \{0 \text{ if } \mathbf{g} \leq 0, +\infty \text{ otherwise}\}, \quad \Psi_h(\mathbf{h}) = \{0 \text{ if } \mathbf{h} = 0, +\infty \text{ otherwise}\},$$

are the so-called indicator functions for the inequality and equality constraints. Then the convex constrained problem (96) can be written in the following canonical form

$$\min \{\Pi(\mathbf{x}) = f(\mathbf{x}) + \Phi_0(\boldsymbol{\Lambda}_0(\mathbf{x})) | \forall \mathbf{x} \in \mathbb{R}\}. \quad (97)$$

By the fact that the canonical function $\Phi_0(\boldsymbol{\xi}_0)$ is convex and lower semi-continuous, the canonical duality relations (10) should be replaced by the following subdifferential forms [55]:

$$\boldsymbol{\xi}_0^* \in \partial\Phi_0(\boldsymbol{\xi}_0) \Leftrightarrow \boldsymbol{\xi}_0 \in \partial\Phi_0^*(\boldsymbol{\xi}_0^*) \Leftrightarrow \Phi_0(\boldsymbol{\xi}_0) + \Phi_0^*(\boldsymbol{\xi}_0^*) = \boldsymbol{\xi}_0^T \boldsymbol{\xi}_0^*, \quad (98)$$

where $\Phi_0^*(\boldsymbol{\xi}_0^*) = \Psi_g^*(\boldsymbol{\lambda}) + \Psi_h^*(\boldsymbol{\mu})$ is the Fenchel conjugate of $\Phi_0(\boldsymbol{\xi}_0)$ and $\boldsymbol{\xi}_0^* = (\boldsymbol{\lambda}, \boldsymbol{\mu})$. By the Fenchel transformation, we have

$$\Psi_g^*(\boldsymbol{\lambda}) = \sup_{\mathbf{g} \in \mathbb{R}^m} \{\mathbf{g}^T \boldsymbol{\lambda} - \Psi_g(\mathbf{g})\} = \begin{cases} 0 & \text{if } \boldsymbol{\lambda} \geq 0 \\ +\infty & \text{otherwise,} \end{cases}$$

$$\Psi_h^*(\boldsymbol{\mu}) = \sup_{\mathbf{h} \in \mathbb{R}^p} \{\mathbf{h}^T \boldsymbol{\mu} - \Psi_h(\mathbf{h})\} = 0 \quad \forall \boldsymbol{\mu} \in \mathbb{R}^p.$$

It is easy to verify that for the indicator $\Psi_g(\mathbf{g})$, the canonical duality leads to

$$\begin{aligned} \boldsymbol{\lambda} \in \partial\Psi_g(\mathbf{g}) &\implies \boldsymbol{\lambda} \geq 0 \\ \mathbf{g} \in \partial\Psi_g^*(\boldsymbol{\lambda}) &\implies \mathbf{g} \leq 0 \\ \boldsymbol{\lambda}^T \mathbf{g} = \Psi_g(\mathbf{g}) + \Psi_g^*(\boldsymbol{\lambda}) &\implies \boldsymbol{\lambda}^T \mathbf{g} = 0, \end{aligned} \quad (99)$$

which are the KKT conditions for the inequality constraints $\mathbf{g}(\mathbf{x}) \leq 0$. While for $\Psi_h(\mathbf{h})$, the canonical duality lead to

$$\begin{aligned} \boldsymbol{\mu} \in \partial\Psi_h(\mathbf{h}) &\implies \boldsymbol{\mu} \in \mathbb{R}^p \\ \mathbf{h} \in \partial\Psi_h^*(\boldsymbol{\mu}) &\implies \mathbf{h} = 0 \\ \boldsymbol{\mu}^T \mathbf{h} = \Psi_h(\mathbf{h}) + \Psi_h^*(\boldsymbol{\mu}) &\implies \boldsymbol{\mu}^T \mathbf{h} = 0. \end{aligned} \quad (100)$$

From the second and third conditions in the (100), it is clear that in order to enforce the constrain $\mathbf{h}(\mathbf{x}) = 0$, the dual variable $\boldsymbol{\mu} = \{\mu_i\}$ must be not zero $\forall i \in I_p$. This is a special complementarity condition for equality constraints, generally not mentioned in many textbooks. However, the implicit constraint $\boldsymbol{\mu} \neq 0$ is important in nonconvex optimization.

By using the Fenchel-Young equality $\Phi_0(\boldsymbol{\xi}_0) = \boldsymbol{\xi}_0^T \boldsymbol{\xi}_0^* - \Phi_0^*(\boldsymbol{\xi}_0^*)$ to replace $\Phi_0(\boldsymbol{\Lambda}_0(\mathbf{x}))$ in (97), the total complementarity function can be obtained in the following form:

$$\Xi_0(\mathbf{x}, \boldsymbol{\xi}_0^*) = f(\mathbf{x}) + [\boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) - \Psi_g^*(\boldsymbol{\lambda})] + [\boldsymbol{\mu}^T \mathbf{h}(\mathbf{x}) - \Psi_h^*(\boldsymbol{\mu})]. \quad (101)$$

Let $\boldsymbol{\sigma}_0 = (\boldsymbol{\lambda}, \boldsymbol{\mu})$. The dual feasible spaces should be defined as

$$\mathcal{S}_0 = \{\boldsymbol{\sigma}_0 = (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}^{m \times p} \mid \lambda_i \geq 0 \ \forall i \in I_m, \ \mu_j \neq 0 \ \forall j \in I_p\}.$$

Thus, on the feasible space $\mathbb{R}^n \times \mathcal{S}_0$, the total complementary function (101) can be simplified as

$$\Xi_0(\mathbf{x}, \boldsymbol{\sigma}_0) = f(\mathbf{x}) + \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x}) + \boldsymbol{\mu}^T \mathbf{h}(\mathbf{x}) = \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}), \quad (102)$$

which is the classical Lagrangian and we have

$$P(\mathbf{x}) = \sup \{\Xi_0(\mathbf{x}, \boldsymbol{\sigma}_0) \mid \forall \boldsymbol{\sigma}_0 \in \mathcal{S}_0\}.$$

This shows that the canonical duality theory is an extension of the Lagrangian theory (indeed, the total complementary function was called the extended Lagrangian in [53]).

For nonconvex constrained problems, the so-called *sequential canonical transformation* (see Chapter 4, [53])

$$\boldsymbol{\Lambda}_0(\boldsymbol{\Lambda}_1(\dots(\boldsymbol{\Lambda}_k(\mathbf{x}))\dots))$$

can be used for target function and constraints to obtain high-order canonical dual problem. Applications have been given to the high-order polynomial optimization [66, 93], nonconvex analysis [53], neural network [106], and nonconvex constrained problems [82, 91, 107, 121, 165].

5.5 SDP relaxation and canonical primal-dual algorithms

Recall the primal problem (\mathcal{P}) (13)

$$(\mathcal{P}) : \min\{\Pi(\boldsymbol{\chi}) = \Phi(\Lambda(D\boldsymbol{\chi})) - \langle \boldsymbol{\chi}, \bar{\boldsymbol{\chi}}^* \rangle \mid \boldsymbol{\chi} \in \mathcal{X}_c\}$$

and its canonical dual (\mathcal{P}^d) (33)

$$(\mathcal{P}^d) : \max \left\{ \Pi^d(\mathbf{S}) = -G_{ap}^*(\mathbf{S}) - \Phi^*(\mathbf{S}) \mid \mathbf{S} \in \mathcal{S}_c^+ \right\},$$

where $G_{ap}^*(\mathbf{S}) = \frac{1}{2} \langle \mathbf{G}^{-1}(\mathbf{S}) \mathbf{F}(\mathbf{S}), \mathbf{F}(\mathbf{S}) \rangle$ is the pure gap function. By the fact that (\mathcal{P}^d) is a concave maximization on a convex domain \mathcal{S}_c^+ , this canonical dual can be solved easily if $\Pi^d(\boldsymbol{\xi}^*)$ has a stationary point in \mathcal{S}_c^+ . For many challenging (NP-hard) problems, the stationary points $\Pi^d(\mathbf{S})$ are usually located on the boundary of $\mathcal{S}_c^+ = \{\mathbf{S} \in \mathcal{S}_c \mid \mathbf{G}(\mathbf{S}) \succ 0\}$. In this case, the matrix $\mathbf{G}(\mathbf{S})$ is singular and the canonical dual problem could have multiple solutions. Two methods can be suggested for solving this challenging case.

1) SDP Relaxation. By using the Schur complement Lemma, the canonical dual problem (\mathcal{P}^d) can be relaxed as [163]

$$(\mathcal{P}^r) : \min \Phi^*(\mathbf{S}) \text{ s.t. } \begin{pmatrix} \mathbf{G}(\mathbf{S}) & \mathbf{F}(\mathbf{S}) \\ \mathbf{F}^T(\mathbf{S}) & 2G_{ap}(\mathbf{S}) \end{pmatrix} \succeq 0, \quad \forall \mathbf{S} \in \mathcal{S}_c. \quad (103)$$

Since $\Phi^*(\mathbf{S})$ is convex and the feasible space is closed, this relaxed canonical dual problem has at least one solution $\bar{\mathbf{S}}$. The associated $\bar{\chi} = \mathbf{G}(\bar{\mathbf{S}})^{-1}\mathbf{F}(\bar{\mathbf{S}})$ is a solution to (\mathcal{P}) only if $\bar{\mathbf{S}}$ is a stationary point of $\Pi^d(\mathbf{S})$. Particularly, if $\Phi^*(\mathbf{S}) = \langle \mathbf{Q}; \mathbf{S} \rangle$ is linear, $\mathbf{F} = 0$, $\mathbf{G}(\mathbf{S}) = \mathbf{S}$, and

$$\mathcal{S}_c = \{\mathbf{S} \in \mathbb{S}_n \mid \langle \mathbf{A}_i; \mathbf{S} \rangle = b_i \quad \forall i \in I_m\}$$

is a linear manifold, where $\mathbb{S}_n = \{\mathbf{S} \in \mathbb{R}^{n \times n} \mid \mathbf{S} = \mathbf{S}^T\}$ is a symmetrical $n \times n$ -matrix space, $\mathbf{Q}, \mathbf{A}_i \in \mathbb{S}_n$ $i \in I_m$ are given matrices and $\mathbf{b} = \{b_i\} \in \mathbb{R}^m$ is a given vector, then by the notation $\mathbf{Q} \bullet \mathbf{S} = \langle \mathbf{Q}; \mathbf{S} \rangle = \text{tr}(\mathbf{Q} \cdot \mathbf{C}) = \mathbf{Q} : \mathbf{C}$, the relaxed canonical dual problem can be written as

$$\min \mathbf{Q} \bullet \mathbf{S} \text{ s.t. } \mathbf{S} \succeq 0, \quad \mathbf{A}_i \bullet \mathbf{S} = b_i, \quad \forall i \in I_m, \quad (104)$$

which is a typical Semi-Definite Programming (SDP) problem in optimization [149]. This shows that the popular SDP problem is indeed a special case of the canonical duality-triality theory for solving the general global optimization problem (3). The SDP method and algorithms have been well-studied in global optimization. But this method provides only a lower bound approach for the global minimal solution to (\mathcal{P}) if its canonical dual has no stationary point in \mathcal{S}_c^+ . Also, in many real-world applications, the local solutions are also important. Therefore, a second method is needed.

2) Quadratic perturbation and canonical primal-dual algorithm. By introducing a quadratic perturbation, the total complementary function (25) can be written as

$$\begin{aligned} \Xi_{\delta_k}(\chi, \xi^*) &= \Xi(\chi, \mathbf{S}) + \frac{1}{2}\delta_k \|\chi - \chi_k\|^2 \\ &= \frac{1}{2}\langle \chi, \mathbf{G}_{\delta_k}(\mathbf{S})\chi \rangle - \Phi^*(\mathbf{S}) - \langle \chi, \mathbf{F}_{\delta_k}(\mathbf{S}) \rangle + \frac{1}{2}\delta_k \langle \chi_k, \chi_k \rangle, \end{aligned}$$

where $\delta_k > 0$, χ_k $k \in I_p$ are perturbation parameters, $\mathbf{G}_{\delta_k}(\mathbf{S}) = \mathbf{G}(\mathbf{S}) + \delta_k \mathbf{I}$, $\mathbf{F}_{\delta_k}(\mathbf{S}) = \mathbf{F}(\mathbf{S}) + \delta_k \chi_k$. Thus, the original canonical dual feasible space \mathcal{S}_c^+ can be enlarged to $\mathcal{S}_{\delta_k}^+ = \{\mathbf{S} \in \mathcal{S}_c \mid \mathbf{G}_{\delta_k}(\mathbf{S}) \succ 0\}$. Using the perturbed total complementary function Ξ_{δ_k} , the perturbed canonical dual problem can be proposed

$$(\mathcal{P}_k^d) : \max \left\{ \min \{ \Xi_{\delta_k}(\chi, \mathbf{S}) \mid \chi \in \mathcal{X}_a \} \mid \mathbf{S} \in \mathcal{S}_{\delta_k}^+ \right\} \quad (105)$$

Based on this perturbed canonical dual problem, a canonical primaldual algorithm has been developed [158, 163].

Canonical Primal-Dual Algorithm. Given initial data $\delta_0 > 0$, $\chi_0 \in \mathcal{X}_a$, and error allowance $\omega > 0$. Let $k = 1$.

- 1) Solve the perturbed canonical dual problem (\mathcal{P}_k^d) to obtain $\mathbf{S}_k \in \mathcal{S}_{\delta_k}^+$.
- 2) Computer $\bar{\chi}_k = [\mathbf{G}_{\delta_k}(\mathbf{S}_k)]^{-1}\mathbf{F}_{\delta_k}(\mathbf{S}_k)$ and let

$$\chi_k = \chi_{k-1} + \beta_k(\bar{\chi}_k - \chi_{k-1}), \quad \beta_k \in [0, 1].$$

- 3) If $|\Pi(\boldsymbol{\chi}_k) - \Pi(\boldsymbol{\chi}_{k-1})| \leq \omega$, then stop, $\boldsymbol{\chi}_k$ is the optimal solution to (\mathcal{P}) .
Otherwise, let $k = k + 1$, go back to 1).

In this algorithm, $\{\beta_k\}$ are given the parameters, which change the search directions. Clearly, if $\beta_k = 1$, we have $\boldsymbol{\chi}_k = \bar{\boldsymbol{\chi}}_k$. This algorithm has been used successfully for solving a class of benchmark problems and sensor network optimization problems [140, 163].

Let $\mathcal{S}_{\delta_k}^- = \{\mathbf{S} \in \mathcal{S}_c \mid \mathbf{G}_{\delta_k}(\mathbf{S}) \prec 0\}$. The combination of this algorithm with the double-min and double-max dualities in the triality theory can be used for finding local optimal solutions [13].

6 Challenges and breakthrough

In the history of sciences, a ground-breaking theory usually has to pass through serious arguments and challenges. This is duality nature and certainly true for the canonical duality-triality theory, which has benefited from recent challenges by M. Voisei, C. Zălinescu and his former student R. Strugariu in a set of 11 papers. These papers fall naturally into three interesting groups.

6.1 Group 1: Bi-level duality

One paper in this group by Voisei and Zălinescu [153] challenges Gao and Yang's work for solving the following minimal distance between two surfaces [91]

$$\min \left\{ \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|^2 \mid g(\mathbf{x}) = 0, h(\mathbf{y}) = 0, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \right\}, \quad (106)$$

where $g(\mathbf{x})$ is convex, while $h(\mathbf{y})$ is a nonconvex function. By the canonical transformation $h(\mathbf{y}) = V(\Lambda(\mathbf{y})) - \mathbf{y}^T \mathbf{f}$, the Gao-Strang complementary function was written in the form of $\Xi(\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}_0, \varsigma)$, where $\boldsymbol{\sigma}_0 = \{\lambda, \mu\}$ is the first level canonical dual variable, i.e. the Lagrange multiplier for $\{g(\mathbf{x}) = 0, h(\mathbf{y}) = 0\}$, while ς is the second level canonical dual variable for the nonconvex constraint (see equation (11) in [91]). Using one counterexample

$$g(\mathbf{x}) = \frac{1}{2}(\|\mathbf{x}\|^2 - 1), \quad h(\mathbf{y}) = \frac{1}{2} \left(\frac{1}{2} \|\mathbf{y} - \mathbf{c}\|^2 - 1 \right)^2 - \mathbf{f}^T(\mathbf{y} - \mathbf{c}), \quad (107)$$

with $n = 2$ and $\mathbf{c} = (1, 0)$, $\mathbf{f} = (\frac{\sqrt{6}}{96}, 0)$, Voisei and Zălinescu proved that “the main results in Gao and Yang [91] are false” and they concluded: “The consideration of the function Ξ is useless, at least for the problem studied in [91]”.

This paper raises up two issues on different levels.

The first issue is elementary: there is indeed a mistake in Gao and Yang's work, i.e. instead of $(\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}_0, \varsigma)$ used in [91], the variables in the total complementary function Ξ should be the vectors $\boldsymbol{\chi} = (\mathbf{x}, \mathbf{y})$ and $(\boldsymbol{\sigma}_0, \varsigma)$ since $\Xi(\boldsymbol{\chi}, \boldsymbol{\sigma}_0, \varsigma)$ is convex in \mathbf{x} and \mathbf{y} but may not in $\boldsymbol{\chi}$. This mistake has been easily corrected in [121]. Therefore, the duality on this level is: opposite to Voisei and Zălinescu's conclusion, the consideration of the Gao-Strang total complementary function Ξ is indeed quite useful for solving the challenging problem (106) [121].

The second issue is crucial. The “counterexample” (107) has two global minimal solutions due to the symmetry (see Fig. 4). Similar to Example 1, the canonical dual

problem (33) $\max\{\Pi^d(\boldsymbol{\sigma}_0, \varsigma) | (\boldsymbol{\sigma}_0, \varsigma) \in \mathcal{S}_c^+\}$ has two stationary points on the boundary of \mathcal{S}_c^+ (cf. Fig. 1(b)). Such case has been discussed by Gao in integer programming problem [68]. It was first realized that many so-called NP-hard problems in global optimization usually have multiple global minimal solutions and a conjecture was proposed in [68], i.e. a global optimization problem is NP-hard if its canonical dual has no stationary point in \mathcal{S}_c^+ . In order to solve such challenging problems, different perturbation methods have been suggested with successful applications in global optimization [80, 140, 142, 155], including a recent paper on solving hard-case of a trust-region subproblem [16]. For this problem, by simply using linear perturbation $\mathbf{f}_k = (\frac{\sqrt{6}}{96}, \frac{1}{k})$ with $|k| \gg 1$, both global minimal solutions can be easily obtained by the canonical duality-triality theory [121] (see Fig. 4 and Fig. 1(a)). Therefore, the duality on this level is: Voisei and Zălinescu's "counterexample" does not contradict the canonical duality-triality theory even in this crucial case.

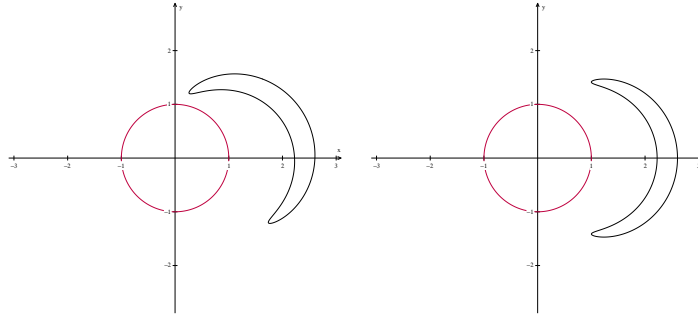


Figure 4: Perturbations for breaking symmetry with $k = 64$ (left) and $k = 10^5$ (right).

Actually, by the general model (3), the nonconvex hyper-surface $h(\mathbf{y})$ in this paper can be written as $h(\mathbf{y}) = W(D\mathbf{y}) - F(\mathbf{y})$, where the double-well function $W(D\mathbf{y})$ is objective (also isotropic), which represents the modeling with symmetry; while the linear term $F(\mathbf{y})$ is a subjective function, which breaks the symmetry and leads to a particular problem. By the fact that nothing in this world is perfect, therefore, any real-world problem must have certain input or defects. This simple truth lays a foundation for the perturbation method and the triality theory for solving challenging problems. However, this fact is not well-recognized in mathematical optimization and computational science⁹, it turns out that many challenges and NP-hard problems are artificially proposed.

6.2 Group 2: Conceptual duality

Of four papers in this group, two were published in pure math journals [147, 151] and two were rejected by applied math journals (*ZAMP* and *Q.J. Mech. Appl. Math*). The paper by Voisei and Zălinescu [151] challenges Gao and Strang's original work on solving the general nonconvex variational problem (3) in finite deformation theory. As we discussed in Section 2.2 that the stored energy $W(\boldsymbol{\epsilon})$ must be objective and can't be linear, the deformation operator Λ should be geometrically admissible in order to have the canonical transformation $W(\boldsymbol{\epsilon}) = \Phi(\Lambda(\boldsymbol{\epsilon}))$, and the external energy $F(\boldsymbol{\chi})$ must be linear such that $\bar{\boldsymbol{\chi}}^* = \partial F(\boldsymbol{\chi})$ is the given input. Oppositely, by listing total six counterexamples, Voisei

⁹Indeed, one authors' paper [140] was first submitted to a computational optimization journal and received such a reviewer's comment: "the authors applied a perturbation, which changed the problem mathematically, ... and I suggest an immediate rejection."

and Zălinescu choose a piecewise linear function $g(u, v) = \{u \text{ (if } v = u^2 \text{)} ; 0 \text{ (otherwise)}\}$ as $\Phi(\xi)$, a parametric function $f(t) = (t, t^2)$ as the geometrically nonlinear operator $\Lambda(t)$ (see Example 3.1 in [151]), and quadratic functions as $F(\chi)$ (see Examples 3.2, 3.4, 3.5 and 3.6 in [151]). While in the rest counterexample (Example 3.3 in [151]), they simply let the external energy $F(u) = 0$ and $\Lambda(u) = u^2 - u$.

Clearly, the piecewise linear function listed by Voisei and Zălinescu is not objective and can't be the stored energy for any real material. Also, both $\Lambda(t)$ and $\Lambda(u)$ are simply not strain measures. Such conceptual mistakes are repeatedly made in their recent papers, say in the paper by Strugariu, Voisei, and Zălinescu (Example 3.3 in [147]), they let $(x(t), y(t)) = A(t) = (\frac{1}{2}t^2, t)$ be the geometrical mapping $\xi(t) = \Lambda(t)$ and, in their notation, $f(x, y) = xy^3(x^2 + (x - y^4)^2)^{-1}$ as the stored energy $\Phi(\xi)$.

For quadratic $F(\chi)$, the input $\bar{\chi}^* = \partial F(\chi)$ depends linearly on the output χ , which is called the *follower force*. In this case, the system is not conservative and the traditional variational methods do not apply. In order to study such nonconservative minimization problems, a so-called rate variational method and duality principle were proposed by Gao and Onat [78]. While for $F(\chi) = 0$, the minimization $\min\{\Pi(\chi) = W(D\chi)\}$ is not a problem but a modelling, which has either trivial solution $\chi = 0$ or multiple solutions $\chi = \text{constant}$ due to certain symmetry of the mathematical modelling. This is a key mistake happened very often in global optimization, which leads to many man-made NP-hard problems as we discussed in the previous subsection.

The concept of a Lagrangian was introduced by J.L. Lagrange in analytic mechanics 1788, which has a standard notation in physics as (see [104])

$$L(\chi) = T(\dot{\chi}) - V(\chi), \quad (108)$$

where T is the kinetic energy and V is the potential energy. By the Legendre transformation $T^*(\mathbf{p}) = \langle \dot{\chi}, \mathbf{p} \rangle - T(\dot{\chi})$, the Lagrangian is also written as

$$L(\chi, \mathbf{p}) = \langle \dot{\chi}, \mathbf{p} \rangle - T^*(\mathbf{p}) - V(\chi). \quad (109)$$

It is commonly known that for problems with linear potential $V(\chi) = \langle \chi, \bar{\chi}^* \rangle$, the Lagrangian $L(\chi)$ is convex and $L(\chi, \mathbf{p})$ is a saddle point functional which leads to a well-known min-max duality in convex systems. But for problems with convex potential $V(\chi)$, the Lagrangian $L(\chi)$ is a d.c. function (difference of convex functions) and $L(\chi, \mathbf{p})$ is not a saddle functional any more. In this case, the Hamiltonian $H(\chi, \mathbf{p}) = \langle \dot{\chi}, \mathbf{p} \rangle - L(\chi, \mathbf{p}) = T^*(\mathbf{p}) + V(\chi)$ is convex. Therefore, a *bi-duality* (i.e. the combination of the double-min and double-max dualities) was proposed in convex Hamilton systems (see Chapter 2 [53]). However, in the paper by Strugariu, Voisei, and Zălinescu [147], the function

$$L(x, y) = \langle a, x \rangle \langle b, y \rangle - \frac{1}{2}\alpha\|x\|^2 - \frac{1}{2}\beta\|y\|^2$$

is defined as the “Lagrangian”, by which, they produced several “counterexamples” for the bi-duality in convex Hamilton systems. In this “Lagrangian”, if we consider $V(x) = \frac{1}{2}\alpha\|x\|^2$ as a potential energy and $T^*(y) = \frac{1}{2}\beta\|y\|^2$ as the complementary kinetic energy, but the term $\langle a, x \rangle \langle b, y \rangle$ is not the bilinear form $\langle Dx; y \rangle$ required in Lagrange mechanics, where D is a differential operator such that Dx and y form a (constitutive) duality pair. This term does not make any sense in Lagrangian mechanics [104] and duality theory [25]. Therefore, the “Lagrangian” used by Strugariu, Voisei, and Zălinescu for producing counterexamples of the bi-duality theory is not the Lagrangian used in

Gao's book [53], i.e. the standard Lagrangian in classical mechanics [104, 133], convex analysis [25], and modern physics [22, 99]. Actually, the bi-duality theory in finite dimensional space is a corollary of the so-called *Iso-Index Theorem* and the proof was given in Gao's book (see Theorem 5.3.6 and Corollary 5.3.1 [53]).

Papers in this group show a big gap between mathematical physics/analysis and optimization. As V.I. Arnold said [3]: "In the middle of the twentieth century it was attempted to divide physics and mathematics. The consequences turned out to be catastrophic."

6.3 Group 3: Anti-Triality

Six papers are in this group on the triality theory. By listing simple counterexamples (cf. e.g. [152]), Voisei and Zălinescu claimed: "a correction of this theory is impossible without falling into trivia"¹⁰. However, even some of these counterexamples are correct, they are not new. This type of counterexamples was first discovered by Gao in 2003 [61, 62], i.e. the double-min duality holds under certain additional constraints (see Remark on page 288 [61] and Remark 1 on page 481 [62]). But neither [61] nor [62] was cited by Voisei and Zălinescu in their papers.

As mentioned in Section 2.4, the triality was proposed originally from post-buckling analysis [46] in "either-or" format since the double-max duality is always true but the double-min duality was proved only in one-dimensional nonconvex analysis [53]. Recently, this double-min duality has been proved first for polynomial optimization [88, 120, 119], and then for general global optimization problems [17, 89]. The certain additional constraints are simply the dimensions of the primal problem and its canonical dual should be the same in order to have strong double-min duality. Otherwise, this double-min duality holds weakly in subspaces with elegant symmetrical forms. Therefore, the triality theory now has been proved in global optimization, which should play important roles for solving NP-hard problems in complex systems.

7 Concluding Remarks and Open Problems

In this article we have discussed the existing gaps between nonconvex analysis/mechanics and global optimization. Common misunderstandings and confusions on some basic concepts have been addressed and clarified, including the objectivity, nonlinearity, and Lagrangian. By the fact that the canonical duality is a fundamental law in nature, the canonical duality-triality theory is indeed powerful for unified understanding complicated phenomena and solving challenging problems. So far, this theory can be summarized for having the following functions:

1. To correctly model complex phenomena in multi-scale systems within a unified framework [53, 61, 92].
2. To solve a large class of nonconvex/nonsmooth/discrete global optimization problems for obtaining both global and local optimal solutions.
3. To reformulate certain nonlinear partial differential equations in algebraic forms with possibility to obtain all possible analytical solutions [54, 70, 76, 75].

¹⁰This sentence is deleted by Voisei and Zălinescu in their revision of [152] after they were informed by referees that their counterexamples are not new and the triality theory has been proved.

4. To understand and identify certain NP-hard problems, i.e., the general global optimization problems are not NP-hard if they can be solved by the canonical duality-triality theory [68, 80, 140].
5. To understand and solve nonlinear (chaotic) dynamic systems by obtaining global stable solutions [139, 115].
6. To check and verify correctness of existing modelling and theories.

There are still many open problems existing in the canonical duality-triality theory. Here we list a few of them.

1. Sufficient condition for the existence of the canonical dual solutions on \mathcal{S}_c^+ .
2. NP-Harness conjecture: A global optimization problem is NP-hard if its canonical dual $\Pi^d(\xi^*)$ has no stationary point on the closed domain $\bar{\mathcal{S}}_c^+ = \{\xi^* \in \mathcal{S}_a \mid \mathbf{G}(\xi^*) \succeq 0\}$.
3. Extremality conditions for stationary points of $\Pi^d(\xi^*)$ on the domain such that $\mathbf{G}(\xi^*)$ is in-definite in order to identify all local extrema.
4. Bi-duality and triality theory for d -dimensional ($d > 1$) nonconvex analysis problems.

The following research topics are challenging:

1. Canonical duality-triality theory for solving bi-level optimization problems.
2. Using least-squares method and canonical duality theory for solving 3-dimensional chaotic dynamical problems, such as Lorenz system and Navier-Stokes equation, etc.
3. Perturbation methods for solving NP-hard integer programming problems, such as quadratic Knapsack problem, TSP, and mixed integer nonlinear programming problems.
4. Unilateral post-buckling problem of the Gao nonlinear beam

$$\min_{\chi \in \mathcal{X}_a} \left\{ \Pi(\chi) = \int_0^L \left[\frac{1}{2} EI \chi_{xx}^2 + \frac{1}{12} \alpha E \chi_x^4 - \frac{1}{2} \lambda E \chi_x^2 - f \chi \right] dx \mid \chi(x) \geq 0 \right\}. \quad (110)$$

Due to the axial compressive load $\lambda > 0$, the downward lateral load $f(x)$ and the unilateral constraint $\chi(x) \geq 0 \quad \forall x \in [0, L]$, the solution of this nonconvex variational problem is a local minimizer of $\Pi(\chi)$ which can be obtained numerically by the canonical dual finite element methods [13, 143] if λ and f are not big enough such that $\chi(x) > 0 \quad \forall x \in [0, L]$. However, if the buckling state $\chi(x) = 0$ happens at any $x \in [0, L]$, the problem could be NP-hard. The open problems include:

- 1) under what conditions for the external loads $\lambda > 0$ and $f(x)$, the problem has a solution $\chi(x) > 0 \quad \forall x \in [0, L]$?
- 2) how to solve the unilateral buckling problem when $\chi(x) = 0$ holds for certain $x \in [0, L]$?

Acknowledgement: This paper is based on a series of plenary lectures presented at international conferences of mathematics, mechanics and global optimization during 2012-2014. Invitations from organizers are sincerely acknowledged. The research has been supported by a grant (AFOSR FA9550-10-1-0487) from the US Air Force Office of Scientific Research. Dr. Ning Ruan was supported by a funding from the Australian Government under the Collaborative Research Networks (CRN) program. The authors sincerely thank Professor D. Steigmann at University of Berkeley for his invitation of editing the special issue of the journal *Math. Mech. Solids* on the canonical duality theory.

References

- [1] Ahn, J, Kuttler, KL, and Shillor, M. (2012). Dynamic contact of two Gao beams, *Electron J Differ Equ.* 194: 1-42.
- [2] Andrews, KT, Dumont, Y, M'Bengue, MF, Purcell, J. and M. Shillor, M. (2012). Analysis and simulations of a nonlinear elastic dynamic beam. *Z Angew Math Phys*, 63: 1005-1019.
- [3] Anorl, VI (1998) On teaching mathematics, *Russian Math. Surveys*, 53 (1), 229 - 236.
- [4] Aspnes, J., Goldberg, D, Yang, Y.R. (2004). On the computational complexity of sensor network localization, in: *Lecture Notes in Computer Science*, 3121, Springer-Verlag, pp. 32-44
- [5] Atiyah, M.F. (2007) Duality in Mathematics and Physics, lecture notes from the Institut de Matematica de la Universitat de Barcelona (IMUB).
- [6] Atai, AA, Steigmann, DJ(1998). Coupled deformations of elastic curves and surfaces. *Int. J. Solids Struct.* 35, 1915-1952.
- [7] Bader B W, Kolda T G (2006). Algorithm 862: MATLAB tensor classes for fast algorithm prototyping. *ACM Trans Math Software*, 32(4): 635-653.
- [8] Ball, J.M. (2002). Some open problems in elasticity. In *Geometry, Mechanics, and Dynamics*, pages 3–59, Springer, New York.
- [9] Bendsoe, M.P. and Sigmund, O. (2003). *Topology Optimization: Theory, Methods and Applications*, Springer, 370 pp.
- [10] Bengue, FM, and Shillor, M. (2008). Regularity result for the problem of vibrations of a nonlinear beam, *Electron J Differ Equ.* 27: 1-12.
- [11] Burachik, R.S., Freire, W.P., Kaya, C.Y. (2014). Interior Epigraph Directions method for nonsmooth and nonconvex optimization via generalized augmented Lagrangian duality, *J. Global Optimization*, 60:501-529.
- [12] Burer, S. and Letchford, A.N. (2012). Non-convex mixed-integer nonlinear programming: A survey. *Surveys in Operations Research and Management Science*, 17, pp. 97-106.
- [13] Cai, K., Gao, DY, Qin, QH (2014). Post-buckling solutions of hyper-elastic beam by canonical dual finite element method, *Mathematics and Mechanics of Solids*. 19(6) 659-671
- [14] Chen, L.L., Gao, D.Y., Ruan, N. (2014). Deterministic method for solving mixed integer quadratic programming problems via polynomial algorithm, working paper.
- [15] Chen, Y. (2014). *Canonical Duality Theory for General Global Optimization Problems and Applications*, PhD. Thesis, School of Science, Information Technology, and Engineering, Federation University Australia.
- [16] Chen Y. and Gao, D.Y. (2014). Global solutions to spherical constrained quadratic minimization via canonical dual approach, <http://arxiv.org/abs/1308.4450>
- [17] Chen Y. and Gao, D.Y. (2014). Global solutions to nonconvex optimization of 4th-order polynomial and log-sum-exp functions, *J. Global Optimization*, DOI 10.1007/s10898-014-0244-5
- [18] Ciarlet, PG (1988). *Mathematical Elasticity*, Volume 1: Three Dimensional Elasticity. North-Holland, 449pp.

- [19] Ciarlet, PG (2013). *Linear and Nonlinear Functional Analysis with Applications*, SIAM, Philadelphia.
- [20] Copeland, E.J. , Kibble, T.W.B. (2010). Cosmic strings and superstrings, *Proc. R. Soc. A*, Vol. 466, no. 2115, 623-657
- [21] Clark, D.M. and Davey, B.A. (1998). *Natural Dualities for the Working Algebraist*, Cambridge University Press.356pp.
- [22] Davier, M (2007). The Higgs boson of the Standard Model, *C. R. Physique*, 8: 986-998
- [23] Dentcheva, D., Römisch, W. (2004). Duality gaps in nonconvex stochastic optimization, *Math Programming*, 101: 515-535.
- [24] Ekeland, I. (2004). Nonconvex duality, in *Proceedings of IUTAM Symposium on Duality, Complementarity and Symmetry in Nonlinear Mechanics*, D.Y. Gao (ed.), pp. 13-19. Kluwer Academic Publishers.
- [25] Ekeland, I. and Temam, R. (1976). *Convex Analysis and Variational Problems*, North-Holland.
- [26] Fang, S.C., Gao, D.Y., Sheu, R.L., and Wu, S.Y.(2008). Canonical dual approach to solving 0-1 quadratic programming problems. *J Industrial and Management Optimization*, 4(1):125
- [27] Fang, S.C. Gao, D.Y., Sheu, R.L., and Xing W. (2009). Global optimization for a class of fractional programming problems. *Journal of Global Optimization*, 45(3):337-353
- [28] Fenchel, W. (1949). On conjugate convex functions, *Canadian J. Math.*, 1:73-77.
- [29] Feng, X. Glowinski, R. and Neilan, M. (2013). Recent developments in numerical methods for fully nonlinear second order partial differential equations, *SIAM Review*, 55(2), pp. 205-267.
- [30] Floudas, C. A. and Gounaris, C. E. (2009). A review of recent advances in global optimization, *J Glob Optim*, 45:3-38.
- [31] Floudas, C.A. and Pardalos, P.M. (2001). *Encyclopedia of Optimization*. Springer, 2-nd edition, 2009.
- [32] Gao, D.Y. (1988). On the complementary bounding theorems for limit analysis, *Int. J. of Solids Struct.*, **24**, 545-556.
- [33] Gao, D.Y. (1988). Panpenalty finite element programming for limit analysis, *Computers & Structures*, 28, pp. 749-755.
- [34] Gao, D.Y. (1990). Dynamically loaded rigid-plastic analysis under large deformation, *Quart. Appl. Math.* **48**, pp. 731-739.
- [35] Gao, D.Y. (1990). On the extremum potential variational principles for geometrical nonlinear thin elastic shell, *Science in China (Scientia Sinica) (A)*, 33 (1), pp. 324-331.
- [36] Gao, D.Y. (1990). On the extremum variational principles for nonlinear elastic plates, *Quart. Appl. Math.*, 48, pp. 361-370.
- [37] Gao, D.Y.(1990). Complementary principles in nonlinear elasticity, *Science in China (Scientia Sinica) (A) (Chinese Ed.)*, 33(4), pp. 386-394.
- [38] Gao, D.Y. (1990). Bounding theorem on finite dynamic deformations of plasticity, *Mech. Research Commun.*, 17, pp. 33-39.

- [39] Gao, D.Y. (1991). Extended bounding theorems for nonlinear limit analysis, *Int. J. Solids Structures*, 27, pp. 523-531.
- [40] Gao, D.Y. (1992). Global extremum criteria for nonlinear elasticity, *J. Appl. Math. Physics (ZAMP)*, 43, 924-937.
- [41] Gao, D.Y. (1995). Duality theory in nonlinear buckling analysis for von Karman equations, *Studies in Appl. Math*, 94, 423-444. MIT
- [42] Gao, D.Y. (1995). Limit analysis of plastic shells subjected to large deflections, *European J. Mech., A/Solids*, 14, no 3, 459-472.
- [43] Gao, D.Y. (1996). Nonlinear elastic beam theory with applications in contact problem and variational approaches, *Mech. Research Commun.*, 23 (1), 11-17.
- [44] Gao, D.Y. (1996). Complementary finite element method for finite deformation nonsmooth mechanics, *J. Eng. Math.*, 30, pp. 339-353.
- [45] Gao, D.Y. (1996). Complementarity and duality in natural sciences, *Philosophical Study in Modern Science and Technology* (in Chinese). Tsinghua University Press, Beijing, China, 12-25. c
- [46] Gao, D.Y. (1997). Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory, *Appl. Mech. Rev.*, 50, 11, November 1997, S64-S71.
- [47] Gao, D.Y. (1998). Duality, triality and complementary extremum principles in nonconvex parametric variational problems with applications, *IMA J. Appl. Math.*, 61, 199-235.
- [48] Gao, D.Y. (1998). Bi-complementarity and duality: A framework in nonlinear equilibria with applications to the contact problems of elastoplastic beam theory, *J. Math. Anal. Appl.*, 221, 672-697.
- [49] Gao, D.Y. (1999). Pure complementary energy principle and triality theory in finite elasticity, *Mech. Res. Comm.* 26 (1), 31-37.
- [50] Gao, D.Y. (1999). *Duality-Mathematics*, Wiley Encyclopedia of Electronical and Electronical Engineering, 6, 68-77.
- [51] Gao, D.Y. (1999). General Analytic Solutions and Complementary Variational Principles for Large Deformation Nonsmooth Mechanics. *Meccanica* 34, 169-198.
- [52] Gao, D. Y.(2000). Finite deformation beam models and triality theory in dynamical post-buckling analysis. *Int. J. Non-Linear Mechanics*, 35, 103-131.
- [53] Gao, D.Y. (2000). *Duality Principles in Nonconvex Systems: Theory, Methods and Applications*, Kluwer Academic Publishers, Dordrecht /Boston /London, xviii + 454pp.
- [54] Gao, D.Y. (2000). Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications, *Nonlinear Analysis*, 42, 7, 1161-1193.
- [55] Gao, D.Y. (2000). Canonical dual transformation method and generalized triality theory in nonsmooth global optimization, *J. Global Optimization*, 17 (1/4), pp. 127-160.
- [56] Gao, D.Y.(2000). Finite deformation beam models and triality theory in dynamical post-buckling analysis, *Int. J. Non-Linear Mechanics*, 5, 103-131.
- [57] Gao, D.Y. (2001). Bi-Duality in Nonconvex Optimization, in *Encyclopedia of Optimization*, C. A. Floudas and P.D. Pardalos (eds). Kluwer Academic Publishers, Dordrecht/Boston/London, Vol. 1, pp. 477-482.

- [58] Gao, D.Y. (2001). Gao, D.Y., Tri-duality in Global Optimization, in *Encyclopedia of Optimization*, C. A. Floudas and P.D. Pardalos (eds). Kluwer Academic Publishers, Dordrecht/Boston/London, Vol. 1, pp. 485-491.
- [59] Gao, D.Y. (2001). Complementarity, polarity and triality in nonsmooth, nonconvex and nonconservative Hamilton systems, *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences*, 359, 2347-2367.
- [60] Gao, D.Y. (2002). Duality and triality in non-smooth, nonconvex and nonconservative systems: A survey, new phenomena and new results, in *Nonsmooth/Nonconvex Mechanics with Applications in Engineering*, edited by C. Baniotopoulos. Thessaloniki, Greece. pp. 1-14.
- [61] Gao, D.Y.(2003). Nonconvex semi-linear problems and canonical dual solutions. *Advances in Mechanics and Mathematics*, Vol. II, D.Y. Gao and R.W. Ogden (ed), Kluwer Academic Publishers, pp. 261-312.
- [62] Gao, D.Y.(2003). Perfect duality theory and complete solutions to a class of global optimization problems. *Optimization*, 52(4-5), 467-493.
- [63] Gao, D.Y.(2004). Canonical duality theory and solutions to constrained nonconvex quadratic programming. *Journal of Global Optimization*, 29, 377-399.
- [64] Gao, D.Y. (2004). Complementary variational principle, algorithm, and complete solutions to phase transitions in solids governed by Landau-Ginzburg equation. *Mathematics and Mechanics of Solid*, 9:285-305.
- [65] Gao, D.Y.(2005). Sufficient conditions and perfect duality in nonconvex minimization with inequality constraints, *J. Industrial and Management Optimization*, 1, 59-69.
- [66] Gao, D.Y.(2006). Complete solutions and extremality criteria to polynomial optimization problems, *J. Global Optimization*, 35, 131-143.
- [67] Gao, D.Y. (2006). Duality in distributed-parameter control of nonconvex and nonconservative dynamical systems with applications. *Nonlinear Dynamics and Systems Theory* 6, 257-279.
- [68] Gao, D.Y. (2007). Solutions and optimality to box constrained nonconvex minimization problems *J. Indust. Manage. Optim.*, 3(2), 293-304.
- [69] Gao, D.Y. (2009). Canonical duality theory: unified understanding and generalized solutions for global optimization. *Comput. & Chem. Eng.* 33, 1964-1972.
- [70] Gao, D.Y. (2014). Analytical solutions to general anti-plane shear problems in finite elasticity, <http://arxiv.org/pdf/1402.6025v2.pdf>
- [71] Gao, D.Y. (2014). Canonical duality-triality: Unified understanding on bifurcation, chaos, and NP-hard problems in complex systems. To appear in *Int. J. Bifurcation and Chaos*.
- [72] Gao, DY and Hajilarov, E. (2014). On analytic solutions to 3-d finite deformation problems governed by St Venant-Kirchhoff material.
- [73] Gao, DY and Lu, X.J. (2014). An optimal mass transfer problem in the limit case, *Math. Meth. Appl. Sci.*
- [74] Gao, D.Y. and Motreanu, D. (2010). *Handbook of Nonconvex Analysis and Applications*. International Press, Boston. 680pp.
- [75] Gao, D.Y., Ogden, R.W. (2008). Multi-solutions to non-convex variational problems with implications for phase transitions and numerical computation. *Q. J. Mech. Appl. Math.* 61, 497-522.

- [76] Gao, D.Y., Ogden, R.W. (2008). Closed-form solutions, extremality and nonsmoothness criteria in a large deformation elasticity problem. *Zeits. Ang. Math. Physik* 59, 498-517.
- [77] Gao, DY, Ogden, RW, Stavroulakis, G (2001). *Nonsmooth and Nonconvex Mechanics: Modelling, Analysis and Numerical Methods*. Kluwer Academic Publishers.
- [78] Gao, D.Y. and Onat, E.T. (1990). Rate variational extremum principles for finite elastoplasticity, *Appl. Math. Mech.*, 11, 7, pp. 659-667.
- [79] Gao, DY and Ruan, N. (2008). Solutions and optimality criteria for nonconvex quadratic-exponential minimization problem, *Math Meth Oper Res*, 67(3):479-491.
- [80] Gao, D.Y., Ruan, N. (2010). Solutions to quadratic minimization problems with box and integer constraints. *J. Glob. Optim.* 47, 463-484.
- [81] Gao, D.Y., Ruan, N., Pardalos, P. (2012). Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization, in: *Sensors: Theory, Algorithms, and Applications*, V. Boginski, C.W. Commander, P. Pardalos, Y. Ye (Eds.), vol. 61, Springer, New York, pp. 37-54.
- [82] Gao, D.Y., Ruan, N., Serali, H. (2009). Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality. *J. Glob. Optim.* 45, 473-497.
- [83] Gao, D.Y., Ruan, N., Serali, H. (2010). Canonical dual solutions for fixed cost quadratic programs, *Optimization and Optimal Control*, A. Chinchuluun et al. (eds.), Springer Optimization and Its Applications 39, DOI 10.1007/978-0-387-89496-6 7.
- [84] Gao, D.Y. and Serali, H.D. (2009). Canonical duality: Connection between nonconvex mechanics and global optimization, in *Advances in Appl. Mathematics and Global Optimization*, 249-316, Springer (2009).
- [85] Gao, D.Y. and Strang, G. (1989). Geometric nonlinearity: Potential energy, complementary energy, and the gap function. *Quart. Appl. Math.*, 47(3), 487-504.
- [86] Gao, D.Y. and Strang, G. (1989). Dual extremum principles in finite deformation elastoplastic analysis, *Acta Appl. Math.*, 17, pp. 257-267.
- [87] Gao, D.Y., L. T. Watson, L.T., Easterling, D. R., Thacker, W.I., and S. C. Billups (2013). Solving the canonical dual of box- and integer-constrained nonconvex quadratic programs via a deterministic direct search algorithm. *Optim. Methods and Software*. 28 (2), pp. 313-326
- [88] Gao, D.Y. and Wu, C.Z. (2012). On the triality theory for a quartic polynomial optimization problem, *J. Ind. Manag. Optim.* 8(1), 229-242.
- [89] Gao, DY, and Wu, C. (2012). On the Triality Theory in Global Optimization, to appear in *J. Global Optimization* <http://arxiv.org/abs/1104.2970>
- [90] Gao, D.Y. and Yang, W.H. (1995). Multi-duality in minimal surface type problems, *Studies in Appl. Math.*, 95, 127-146. MIT.
- [91] Gao, D.Y. and Yang, W.-C. (2008), Complete solutions to minimal distance problem between two nonconvex surfaces, *Optimization*, 57(5), 705-714.
- [92] Gao, D.Y., Yu, H.F. (2008). Multi-scale modelling and canonical dual finite element method in phase transitions of solids. *Int. J. Solids Struct.* 45, 3660-3673.

- [93] Gao, T.K. (2013). Complete solutions to a class of eighth-order polynomial optimization problems, *IMA J Appl Math*, published online July 21, 2013 doi:10.1093/imamat/hxt033
- [94] Grossmann, I.E. (2002). Review of nonlinear mixed-integer and disjunctive programming techniques, *Optimization and Engineering*, 3, 227-252.
- [95] Hellinger, E. (1914). Die allgemeine Ansätze der Mechanik der Kontinua. *Encyklopädie der Mathematischen Wissenschaften* IV, 4, pp. 602-94.
- [96] Holzapfel, G.A. (2000). *Nonlinear Solid Mechanics: A Continuum Approach for Engineering*. Wiley. ISBN 978-0471823193.
- [97] Horst, R., Pardalos, P.M., and Thoai, N.V.(2000). *Introduction to Global Optimization*. Kluwer Academic Publishers.
- [98] Ionitǎ, A. (2001). *Finite Element Analysis of the Deformation of a Rubber Diaphragm*, PhD Thesis, Department of Engineering Science and Mechanics, Virginia Tech. 142pp.
- [99] Jaffe, A. and Witten, E. (2006). Quantum Yang-Mills Theory, in *The millennium prize problems*, by J. Carlson, A. Jaffe and A. Wiles (eds). Clay Mathematics Institute. pp. 129-152.
- [100] Karp, R. (1972). Reducibility among combinatorial problems. In: Miller, R.E., Thatcher, J.W. (eds.) *Complexity of Computer Computations*, pp. 85-103. Plenum Press, New York.
- [101] Koiter, W.T. (1976). On the complementary energy theorem in nonlinear elasticity theory, in *Trends in the Application of Pure Mathematics to Mechanics*, G. Fichera (ed.). Pitman.
- [102] Kuttler, KL, Purcell, J, and Shillor, M. (2012). Analysis and simulations of a contact problem for a nonlinear dynamic beam with a crack. *Q J Mech Appl Math*. 65: 1-25.
- [103] Lagrange, J. L. (1788). *Méchanique analytique*, La veuve Desaint, Paris; 2nd ed., *Mecanique analytique*, Courcier, Paris, 1811 - 1815 (2 vols.)
- [104] Landau, L.D. and Lifshitz, E.M. (1976). *Mechanics*. Vol. 1 (3rd ed.). Butterworth-Heinemann. ISBN 978-0-750-62896-9.
- [105] Laozi, *Dao De Jing*, 500-400 B.C. See <http://en.wikipedia.org/wiki/Laozi>
- [106] Latorre, V. and Gao, D.Y. (2014). Canonical duality for RBF neural networks, *Neuralcomputings*, 134: 189-197.
- [107] Latorre, V. and Gao, D.Y. (2014). Canonical duality for solving general nonconvex constrained problems, to appear in *Optimization Letters*.
- [108] Latorre, V. and Sagratella, S. (2014). A Canonical Duality Approach for the Solution of Affine Quasi-Variational Inequalities. *J. of Global Optim.* doi: 10.1007/s10898-014-0236-5.
- [109] Latorre, V., Ruan, N., Gao, D.Y. (2014). Unified canonical duality methodology for global optimization. To appear *Mathematics of Mechanics and Solids*.
- [110] Latorre, V. (2014). A Potential Reduction Method for Canonical Duality, with an Application to the Sensor Network Localization Problem. Submitted to *J. of Global Optim.* Preprint available at <http://arxiv.org/pdf/1403.5991.pdf>.
- [111] Lee, S. J. and Shield, R. T. (1980). Variational principles in finite elastics, *J. Appl. Math. Physics (ZAMP)*, 31, 437-453.
- [112] Levinson, M. (1965). The complementary energy theorem in finite elasticity, *Trans. ASME, ser. E, J. Appl. Mech.*, 87, pp. 826-828.

- [113] Li, J., Kuttler, K.L., Shillor, M. (2014). Existence for dynamic contact of a stochastic viscoelastic Gao beam, *Nonlinear Analysis, Series B: Real World Applications*.
- [114] Li, S.F. and Gupta, A.(2006). On dual configuration forces, *J. of Elasticity*, 84, 13-31.
- [115] Li, C., Zhou, X., and Gao DY (2014). Stable trajectory of logistic map, *Nonlinear Dynamics*, DOI 10.1007/s11071-014-1433-y
- [116] Liu, I.-S. (2005). Further remarks on Euclidean objectivity and the principle of material frame-indifference. *Continuum Mech. Thermodyn.*, 17, 125-133
- [117] Machalova, J, and Netuka, H. (2011). Bending of a nonlinear beam reposing on an unilateral foundation, *Appl Comput Mech.* 5: 45-54.
- [118] Marsden, J.E. and Hughes, T.J.R.(1983). *Mathematical Foundations of Elasticity*, Prentice-Hall, 1983.
- [119] Morales, D.M. and Gao, D.Y. (2014). Canonical duality theory and triality for solving general nonconstrained global optimization problems, to appear in *Mathematics and Mechanics of Complex Systems*. <http://arxiv.org/abs/1210.0180>.
- [120] Morales-Silva, D M and Gao, D Y (2013). Complete solutions and triality theory to a nonconvex optimization problem with double-well potential in R^n , *Numer. Algebra Contr. Optim.*, 3 (2): 271-282.
- [121] Morales-Silva, D. and Gao, D.Y. (2014). On the minimal distance between two surfaces, *J. Glob. Opt.*, published online at <http://arxiv.org/abs/1210.1618>
- [122] Moreau, J.J. (1962). Fonctions convexes en dualité, *Séminaires de Mathématiques*, Faculté des Sciences, Université de Montpellier.
- [123] Moreau, J.J. (1968). La notion de sur-potentiel et les liaisons unilatérales en élastostatique. *C.R. Acad. Sci. Paris* 267 A, 954-957.
- [124] Moreau, J.J., Panagiotopoulos, P.D., Strang, G. (1988). *Topics in Nonsmooth Mechanics*. Birkhauser.
- [125] Morris, S.A. (1977). *Pontryagin Duality and Structure of Locally Compact Abelian Groups*. Cambridge University Press. 128pp.
- [126] Murdoch, A.I.(2003). Objectivity in classical continuum physics: a rationale for discarding the principle of invariance under superposed rigid body motions in favour of purely objective considerations, *Continuum Mech. Thermodyn.*, 15: 309-320.
- [127] Murdoch, A.I.(2005). On criticism of the nature of objectivity in classical continuum physics, *Continuum Mech. Thermodyn.*, 17(2):135-148
- [128] Oden, J.T. and Reddy, J.N. (1973). On dual complementary variational principles in mathematical physics, *Int. J. Engrg. Sci.* 12, 1-29.
- [129] Oden, J.T. and Reddy, J.N. (1983). *Variational Methods in Theoretical Mechanics*. Springer-Verlag.
- [130] Ogden, R.W. (1975). A note on variational theorems in non-linear elastostatics. *Math. Proc. Camb. Phil. Soc.* 77, 609-615.
- [131] Ogden, RW (1984/97). *Non-Linear Elastic Deformations*, Ellis Horwood/Dover.
- [132] Pardalos, P.M. and Vavasis, S.A. (1991). Quadratic programming with one negative eigenvalue is NP-hard, *J. Global Optim.* 1, 15-22.

- [133] Panza, M. (2003). The Origins of Analytic Mechanics in the 18th Century, in Hans Niels Jahnke (editor), *A History of Analysis*, p. 137 - 153
- [134] Pappalardo, M. (1986) On the duality gap in nonconvex optimization, *Mathematics of Operations Research*, 11: 30-35.
- [135] Pardalos, P.M., Romeijn, H.E., Tuy, H.(2000). Recent developments and trends in global optimization. *J. Comput. Appl. Math.* 124(1-2), 209-228.
- [136] Raoult, A. (2010). Quasiconvex envelopes in nonlinear elasticity. In *Rank-One Convexity in Applied Mechanics*, J. Schröder and P. Neff (Eds), Springer, pp.17-52.
- [137] Reissner, E (1953), On a variational theorem for finite elastic deformations, *J. Math. & Phys.* 32, 129-153.
- [138] Ruan, N. (2013). *Global Optimization for Nonconvex Optimization Problems*, PhD Thesis, Department of Mathematics and Statistics, Curtin University, 125pp.
- [139] Ruan, N. and Gao, D.Y.(2014). Canonical duality approach for nonlinear dynamical systems, *IMA J. Appl. Math.*, 79: 313-325.
- [140] Ruan, N. and Gao, D.Y. (2014). Global optimal solutions to a general sensor network localization problem, *Performance Evaluations*, 75-76: 1-16.
- [141] Ruan, N. and Gao, D.Y. (2014). Global optimal solution to quadratic discrete programming problem with inequality constraints, to appear in *J. Global Optimization*.
- [142] Ruan, N., Gao, D.Y. and Jiao, Y.(2008). Canonical dual least square method for solving general nonlinear systems of quadratic equations, *Comput. Optim. Appl.*, 47, 335-347.
- [143] Santos, H.A.F.A. and Gao D.Y. (2011) Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam, *Int. J. Nonlinear Mechanics*, 47: 240 - 247. doi:10.1016/j.ijnonlinmec.2011.05.012
- [144] Saxe, J. (1979). Embeddability of weighted graphs in k-space is strongly NP-hard, in *Proc. 17th Allerton Conference in Communications, Control, and Computing*, Monticello, IL, pp.480-489.
- [145] Strang, G. (1986). *Introduction to Applied Mathematics*, Wellesley-Cambridge Press.
- [146] Stumpf, H. (1979). The derivation of the dual extremum and complementary stationary principles in geometrical nonlinear shell theory, *Ing. Arch.* 48, 221-237.
- [147] Strugariu, R., Voisei, M.D., and Zălinescu, C. (2011). Counter-examples in bi-duality, triality and tri-duality. *Discrete and Continuous Dynamical Systems - Series A*, Volume 31, Number 4, December 2011.
- [148] Su, K, Tu, SK, Gao, DY and Xu, L (2009). Canonical dual approach to binary factor analysis, T. Adali et al. (Eds.): ICA 2009, LNCS 5441, pp. 346-353. Springer-Verlag, Berlin, Heidelberg.
- [149] Todd, M. (2001). Semidefinit optimization, *Acta Numerica*, 10:515-560.
- [150] Vavasis, S. (1990). Quadratic programming is in NP, *Info. Proc. Lett.* 36, 73-77.
- [151] Voisei, M.D. and Zălinescu, C.(2011). Some remarks concerning Gao-Strang's complementary gap function, *Applicable Analysis*, Vol. 90, No. 6, 1111-1121
- [152] Voisei, M.D. and Zălinescu, C. (2011). Counterexamples to some triality and tri-duality results. *J. Global Optimization*, 49:173-183, 2011.

- [153] Voisei, M.D. and Zălinescu, C. (2011). A counter-example to “Minimal distance between two nonconvex surfaces”, *Optimization*, 60(5), 593-602.
- [154] Wang, Z.B., Fang, S.C., Gao, D.Y. and Xing, W.X. (2008). Global extremal conditions for multi-integer quadratic programming. *J Industrial and Management Optimization*, 4(2):213.
- [155] Wang, Z.B., Fang, S.C., Gao, D.Y. and Xing, W.X. (2012). Canonical dual approach to solving the maximum cut problem, *J. Glob. Optim.*, 54, 341-351. DOI 10.1007/s10898-012-9881-8
- [156] Wright, S.J. (1997). *Primal-Dual Interior-Point Methods*, SIAM, Philadelphia.
- [157] Wu, C., Gao, D.Y. and Teo, K.L. (2013). A direct optimization method for low group delay FIR filter design, *Signal Processing*, 93(7), pp. 1764-1772.
- [158] Wu, C., Li, C., and Gao, D.Y. (2013). Canonical primal-dual method for solving non-convex minimization problems, <http://arxiv.org/pdf/1212.6492v1.pdf>
- [159] Yau, S.-T. and Gao, D.Y. (1992). Obstacle problems for von Kármán equations, *Adv. Appl. Math.*, 13, pp. 123-141.
- [160] Yu, H.F. (2011). *A Numerical Investigation Of The Canonical Duality Method For Non-Convex Variational Problems*, PhD. Thesis, Department of Mathematics, Virginia Tech., 104pp.
- [161] Zhang J., Gao, D.Y. and Yearwood, J.: A novel canonical dual computational approach for prion AGAAAAGA amyloid fibril molecular modeling. *Journal of Theoretical Biology*, 284, 149-157 (2011). doi:10.1016/j.jtbi.2011.06.024
- [162] Zhou, X.J. (2014). *Canonical Dual Algorithms for Global Optimization with Applications*. PhD Thesis, Federation University Australia, 171pp.
- [163] Zhou, X.J., Gao, D.Y. and Yang, C.H. (2014). Canonical primal-dual algorithm for solving fourth-order polynomial minimization problems, *Applied Mathematics and Computation*, 227: 246-255.
- [164] Zhou, X.J., Gao, D.Y. and Yang, C.H. (2014). Canonical dual approach to continuous time constrained optimal control, *J of Global Optimization*.
- [165] Zhou, X.J., Gao, D.Y. and Yang, C.H. (2014). Global solutions to a class of CEC benchmark constrained optimization problems, *Optimization Letters*, DOI: 10.1007/s11590-014-0784-0